

Braun

E 1357 E

ARCHIV DER ELEKTRISCHEN ÜBERTRAGUNG

Unter Mitarbeit von

H. Busch, R. Feldtkeller, E. Hölzler, K. Hoffmann, O. Kirchner
E. Kramar, A. Mehli, W. Nestel, H. Piloty, L. Pungs, F. Schröter, W. O. Schumann
M. Seddig, A. Siemens, R. Tamm, R. Theile, F. Tank, R. Vieweg, G. Wuckel

Begründet von

KARL WILLY WAGNER

Schriftleitung:

J. Piesch, J. Schunack, M. Strutt, A. Thoma

Geschäftsführende Redaktion:

F. Rühmann

BAND 14

SEPTEMBER 1960

HEFT 9

S. HIRZEL VERLAG · STUTTGART

KWR

KABELWERK RHEYDT AKT.-GES.

RHEYDT (RHLD.)

A. E. Ü. Band 14 [1960], Heft 9, Seiten 367 – 420

INHALT

Originalmitteilungen:

	Seite
Die Ersatzwellenquelle	367
Von <i>H.-J. Butterweck</i>	

Aktive Hochfrequenzspektrometer für ionosphärische Echolotung.	
II. Selektion der Echozeiten aus dem Störpegel	373
Von <i>K. Rawer</i>	

Oberflächenströme in Inversionsschichten an Halbleitern	380
Von <i>E. Groschwitz, E. Hofmeister und R. Ebhardt</i>	

Berechnung der Verzerrungs- und Störeffekte bei Transistor-Verstärkerstufen auf Grund des Ersatzschaltbildes	397
Von <i>J. S. Vogel und M. J. O. Strutt</i>	

Der Satz von den kleinen Kennlinienänderungen und seine Anwendung auf nichtlineare und lineare Netzwerke	405
Von <i>E. Schwartz</i>	

Die Steilheit des HF-Legierungs- und Drifttransistors als Funktion der Frequenz und des Arbeitspunktes und ihre Ableitungen	411
Von <i>W. Minner</i>	

Buchbesprechungen	372, 379, 410
-----------------------------	---------------



Wir suchen für unser Laboratorium für Halbleiter

HF-PHYSIKER oder ELEKTRONIKER

(Dr.-Ing. oder Dipl.-Ing.) der Fachrichtung HF-Technik, Schwachstromtechnik, Elektronik.

Die Tätigkeit erstreckt sich insbesondere auf das Gebiet der Anwendung und Auslegung von Halbleiter-Bauelementen für die HF-Technik.

Es besteht bereits eine Gruppe von Fachkräften, die sich mit den Meßgeräten und Aufbauten auf diesem Gebiet befafzt. Wir suchen einen Herrn für die Leitung und Koordinierung der Arbeiten auf diesem Gebiet. Es wird Gelegenheit zur Einarbeitung in dieses modernste und zukunftsreiche Gebiet geboten, das die Erfassung der elektrischen Kenngrößen von neuen Halbleiterelementen, den Aufbau von Meßgeräten, die Festlegung von Halbleitereigenschaften und den Bau von Apparaten für die Anwendungen modernster Bauelemente (Tunneldioden, Drift-Transistoren, Mesa-Transistoren usw.) überstreicht.

Bewerber, die Interesse an einer verantwortlichen und zukunftsreichen Stellung haben und die Erfahrungen in der Leitung von Fachkräften (Menschenführung) besitzen, bitten wir um Einsendung ihrer Unterlagen (handgeschr. Lebenslauf, Lichtbild, Zeugnisabschriften und Angaben über Gehaltsansprüche und frühesten Eintrittstermin).

Wir sind bekannt für eine gute Bezahlung, für ein gutes Betriebsklima und moderne Sozialeinrichtungen (zusätzliche Altersversorgung, Werkküche und Kantine). Bei der Wohnungsbeschaffung sind wir behilflich.

**Süddeutsche Telefon-Apparate-, Kabel- und Drahtwerke A.G.,
TE-KA-DE, Nürnberg, Nornenstraße 33**

TRÄGERFREQUENZSYSTEM

FGF-V 120

**Zur Übertragung
von 120 Sprechkanälen
je Doppel-Ader in einem
Übertragungsbereich
von 12 bis 252 kHz.**

Die Übertragungstechnischen
Eigenschaften entsprechen
den Empfehlungen des CCITT.



FELTEN & GUILLEAUME FERNMELDEANLAGEN GMBH · NÜRNBERG

ARCHIV

DER ELEKTRISCHEN ÜBERTRAGUNG (A.E.Ü.)

Unter Mitarbeit von H. Busch, R. Feldtkeller, E. Hölzler, K. Hoffmann, O. Kirchner, E. Kramar, A. Mehli, W. Nestel, H. Piloty, L. Pungs, F. Schröter, W. O. Schumann, M. Seddig, A. Siemens, R. Tamm, F. Tank, R. Theile, R. Vieweg, G. Wuckel

Begründet von KARL WILLY WAGNER

Schriftleitung: J. Piesch, J. Schunack, M. Strutt, A. Thoma / Geschäftsführende Redaktion: F. Rühmann
S. HIRZEL VERLAG, Stuttgart

Band 14

September 1960

Heft 9

Die Ersatzwellenquelle

VON HANS-JÜRGEN BUTTERWECK

Mitteilung aus dem Institut für Hochfrequenztechnik der Technischen Hochschule Aachen

(A.E.Ü. 14 [1960], 367 – 372; eingegangen am 13. Juli 1960)

DK 621.372.4

In letzter Zeit geht man häufig dazu über, den Zustand elektrischer Schaltungen statt durch Spannungen und Ströme durch Wellen zu beschreiben. Die elektrischen Eigenschaften der Bauelemente kennzeichnet man in dieser Betrachtungsweise durch ihren Reflexionsfaktor (beim Zweipol) bzw. ihre Streumatrix (beim $2n$ -Pol). Im vorliegenden Aufsatz wird gezeigt, wie sich auch die Zweipol-Quellen (Generatoren) durch Wellengrößen beschreiben lassen. Es wird eine Ersatzwellenquelle hergeleitet, die äquivalent zur bekannten Ersatzspannungs- und Ersatzstromquelle ist. An einigen Beispielen wird die Anwendung der Ersatzwellenquelle gezeigt.

Recently one often proceeds to describing the condition of electrical circuits in terms of waves rather than voltages and currents. In this formulation the electrical characteristics of the components are characterized by their reflection coefficients (in the case of two-terminal networks) and scatter matrices (in that of $2n$ -terminal networks). This paper shows, how wave parameters can also describe two-terminal sources (generators). An equivalent wave source is derived as the counterpart of the well-known equivalent voltage and current source. Its application is shown by reference to some examples.

Vorbemerkung

Wenn im folgenden von Zweipolen die Rede ist, so soll darunter (gemäß Normblatt DIN 1323 [6]) ein elektrisches Gebilde mit zwei zugänglichen Punkten, den sogenannten „Klemmen“, verstanden werden¹. Die zu betrachtenden Zweipole sollen stets linear sein, d. h. zwischen der Klemmenspannung und dem Klemmenstrom soll eine lineare Beziehung bestehen. Wenn der Zweipol keine Leerlaufspannung und keinen Kurzschlußstrom aufweist, wollen wir von einem „quellenlosen Zweipol“ sprechen, anderenfalls von einer „Zweipolquelle“². Es sei darauf hingewiesen, daß der „quellenlose Zweipol“ nicht stets ein passiver Zweipol ist. Z. B. ist ein negativer Widerstand ein quellenloser, aber stets aktiver Zweipol. Umgekehrt ist eine Zweipolquelle nicht stets ein aktiver Zweipol, denn sie kann z. B. bei Zusammenschaltung mit einer anderen Zweipolquelle gegebenenfalls auch Wirkleistung aufnehmen.

¹ In der Mikrowellentechnik ist ein Zweipol (neuerdings auch als „Eintor“ bezeichnet) ein mit einem Hohlleitungsanschluß versehenes Gebilde, dessen Eigenschaften man in einer bestimmten Querschnittsebene des Hohlleiters studiert.

² J. WALLÖT [1] unterscheidet zwischen „Zweipol“ und „Zweipolquelle“, benutzt also keinen besonderen Ausdruck für den Oberbegriff.

1. Ableitung der Ersatzwellenquelle

Bei einer Wahl der Bezugspfeile gemäß Bild 1a („Verbraucherpfeile“) gilt für jeden quellenlosen Zweipol die Gleichung

$$\underline{U} = \underline{Z} \underline{I}, \quad (1)$$

d. h. Klemmenspannung \underline{U} und Klemmenstrom \underline{I} sind zueinander proportional. Der Proportionalitätsfaktor \underline{Z} ist der komplexe Widerstand des Zweipols.

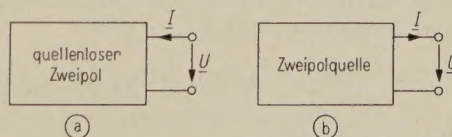


Bild 1. Zweckmäßige Bezugspfeile beim quellenlosen Zweipol (a) und bei der Zweipolquelle (b).

Für jede Zweipolquelle gilt dagegen mit den Zählpfeilen gemäß Bild 1b („Erzeugerpfeile“) nach einem Satz von H. HELMHOLTZ [3] die Gleichung

$$\underline{U} = \underline{U}_q - \underline{Z}_i \underline{I}, \quad (2)$$

d. h. \underline{U} und \underline{I} sind durch eine allgemeine lineare Beziehung miteinander verknüpft. Die Konstante \underline{U}_q

ist die „Leerlaufspannung“, „Urspannung“ oder „Quellenspannung“ der Zweipolquelle, \underline{Z}_i ist ihr komplexer Innenwiderstand. Der Gl. (2) entspricht eine Ersatzspannungsquelle gemäß Bild 2a.

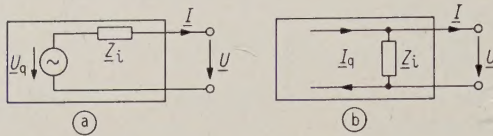


Bild 2. Ersatzspannungsquelle (a) und Ersatzstromquelle (b).

Löst man Gl. (2) nach \underline{I} auf, so erhält man

$$\underline{I} = \frac{\underline{U}_q}{\underline{Z}_i} - \frac{\underline{U}}{\underline{Z}_i} = \underline{I}_q - \frac{\underline{U}}{\underline{Z}_i}. \quad (3)$$

$\underline{I}_q = \underline{U}_q / \underline{Z}_i$ ist der „Kurzschlußstrom“, „Urstrom“ oder „Quellenstrom“. Der Gl. (3) entspricht die von H. F. MAYER [4] angegebene Ersatzstromquelle (Bild 2b).

Für viele Probleme (z. B. bei Leitungen, Filtern, Mikrowellenbauteilen) ist es zweckmäßig, statt mit Spannungen und Strömen mit Wellen zu rechnen. Bezeichnen wir gemäß [2] die auf ein Klemmenpaar zulaufende normierte Spannungswelle mit \underline{a} und die vom Klemmenpaar ablaufende Welle mit \underline{b} (Bild 3), so ergeben sich bei den Verbraucherpeilen gemäß Bild 1a Spannung und Strom an den Klemmen zu

$$\underline{U} = \sqrt{Z_L} (\underline{a} + \underline{b}), \quad \underline{I} = \frac{1}{\sqrt{Z_L}} (\underline{a} - \underline{b}) \quad (4)$$

und bei Verwendung von Erzeugerpeilen gemäß Bild 1b zu

$$\underline{U} = \sqrt{Z_L} (\underline{a} + \underline{b}), \quad \underline{I} = \frac{1}{\sqrt{Z_L}} (\underline{b} - \underline{a}). \quad (5)$$

Z_L ist der reelle Normierungswiderstand, der in der Regel gleich dem Wellenwiderstand der angeschlossenen Leitung ist. \underline{a} und \underline{b} sind so normiert, daß die

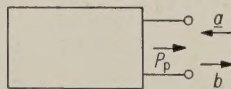


Bild 3. Zu- und ablaufende Wellen beim Zweipol; Bezugspfeil für die Wirkleistung.

von den Klemmen wegströmende Wirkleistung P_p (siehe Bild 3) nach der einfachen Formel

$$P_p = \frac{1}{2} (|\underline{b}|^2 - |\underline{a}|^2) = \frac{1}{2} \operatorname{Re} (\underline{U} \underline{I}^*) \quad (6)$$

berechnet werden kann.

Anmerkungen zu Gl. (6):

- $|\underline{U}|$ und $|\underline{I}|$ sind die Scheitelwerte von Spannung und Strom.
- Es wurde die von den Klemmen wegströmende und nicht die zuströmende Leistung berechnet, weil wir uns im weiteren vorwiegend mit der Zweipolquelle beschäftigen wollen, bei der in Zusammenschaltung mit einem passiven Verbraucher die Leistung tatsächlich wegströmt, also $P_p > 0$ wird.
- Der rechte Teil von Gl. (6) gilt nur für Erzeugerpeile.

Setzt man die Gl. (4) in die Gl. (1) für den *quellenlosen Zweipol* ein, so erhält man das bekannte Ergebnis

$$\underline{b} = \frac{\underline{Z} - \underline{Z}_L}{\underline{Z} + \underline{Z}_L} \underline{a} = \underline{r} \underline{a}, \quad (7)$$

d. h. ablaufende Welle \underline{b} und zulaufende Welle \underline{a} sind zueinander proportional. Der Proportionalitätsfaktor \underline{r} ist der komplexe Reflexionsfaktor des Zweipols.

Setzt man entsprechend die Gl. (5) in die Gl. (2) für die *Zweipolquelle* ein, so erhält man das Ergebnis

$$\underline{b} = \frac{\underline{U}_q}{\sqrt{Z_L} \left(1 + \frac{\underline{Z}_i}{\underline{Z}_L}\right)} + \underline{a} \frac{\underline{Z}_i - \underline{Z}_L}{\underline{Z}_i + \underline{Z}_L},$$

$$\underline{b} = \underline{b}_q + \underline{r}_i \underline{a} \quad (8)$$

mit $\underline{b}_q = \frac{\underline{U}_q}{\sqrt{Z_L} \left(1 + \frac{\underline{Z}_i}{\underline{Z}_L}\right)}, \quad \underline{r}_i = \frac{\underline{Z}_i - \underline{Z}_L}{\underline{Z}_i + \underline{Z}_L};$

d. h. \underline{b} und \underline{a} sind durch eine allgemeine lineare Beziehung miteinander verknüpft. Die Konstante \underline{b}_q ist die „Anpassungswelle“, „Urwelle“ oder „Quellenwelle“ der Zweipolquelle. Sie ist gerade die bei Zusammenschaltung mit einem nichtreflektierenden Verbraucher ($\underline{a} = 0$) von der Quelle ablaufende Welle \underline{b} . \underline{r}_i ist der komplexe, innere Reflexionsfaktor der Quelle. Gl. (8) ist die Grundgleichung der linearen Zweipolquelle in der Darstellung durch Wellengrößen. Sie führt zu dem Satz von der *Ersatzwellenquelle*:

Bei jeder linearen Zweipolquelle ist die von den Klemmen ablaufende Welle \underline{b} die Summe aus einer konstanten Urwelle \underline{b}_q und einer Welle $\underline{r}_i \underline{a}$, die durch Reflexion der zulaufenden Welle \underline{a} am inneren Reflexionsfaktor \underline{r}_i der Quelle entsteht (Symbol der Ersatzwellenquelle Bild 4).

Ebenso wie beim *quellenlosen Zweipol* der Reflexionsfaktor \underline{r} vom Normierungswiderstand abhängt, hängen auch die Elemente \underline{b}_q und \underline{r}_i der Ersatzwellenquelle vom Normierungswiderstand Z_L ab.

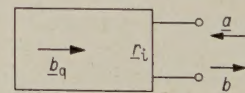


Bild 4. Symbol der Ersatzwellenquelle.

2. Zusammenschaltung einer Zweipolquelle mit einem *quellenlosen Zweipol*

Wir schalten nun gemäß Bild 5 eine durch ihre Ersatzwellenquelle (\underline{b}_q , \underline{r}_i) dargestellte Zweipolquelle mit einem *quellenlosen Zweipol* (gekennzeichnet durch seinen Reflexionsfaktor \underline{r}_a) zusammen.

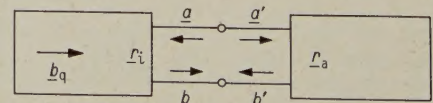


Bild 5. Zusammenschaltung einer Zweipolquelle mit einem *quellenlosen Zweipol*.

Die von der Quelle ablaufende Welle ist die auf \underline{r}_a zulaufende Welle bzw. umgekehrt:

$$\underline{a} = \underline{b}', \quad \underline{b} = \underline{a}'. \quad (9)$$

Außerdem gilt

$$\underline{b} = \underline{b}_q + \underline{r}_i \underline{a}, \quad (10)$$

und mit $\underline{b}' = \underline{r}_a \underline{a}'$ (Gl. (7)) und Gl. (9)

$$\underline{a} = \underline{r}_a \underline{b}. \quad (11)$$

Durch Einsetzen von Gl. (11) in Gl. (10) erhalten wir die von der Quelle ablaufende Welle \underline{b} zu

$$\underline{b} = \frac{\underline{b}_q}{1 - \underline{r}_i \underline{r}_a}. \quad (12)$$

\underline{b} ist also vom inneren Reflexionsfaktor \underline{r}_i und vom äußeren Reflexionsfaktor \underline{r}_a abhängig und ist gleich der Urwelle \underline{b}_q , wenn einer der beiden Reflexionsfaktoren verschwindet. Wenn $\underline{r}_i = 0$ ist, ist also die von der Quelle ablaufende Welle \underline{b} für jeden Abschluß gleich der Urwelle \underline{b}_q .

Wir betrachten des weiteren nur noch die Fälle $|\underline{r}_a| \leq 1$ und $|\underline{r}_i| \leq 1$, d. h. der quellenlose Zweipol sei passiv und der Betrag der an \underline{r}_i reflektierten Welle sei nicht größer als der der zulaufenden Welle. Wegen der Eindeutigkeit des Leistungsflusses können wir dann die Bezeichnungen „Zweipolquelle“ und „quellenloser Zweipol“ durch die geläufigeren Bezeichnungen „Generator“ und „Verbraucher“ ersetzen.

Unter diesen Voraussetzungen kann der Nenner in Gl. (12) $1 - \underline{r}_i \underline{r}_a$ bei vorgegebenem \underline{r}_i und variablem Verbraucher \underline{r}_a in der komplexen Ebene nur noch Werte innerhalb des in Bild 6 schraffierten

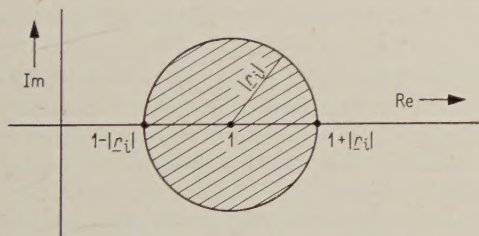


Bild 6. Zur Ableitung der Welle minimaler und maximaler Amplitude.

Kreises annehmen³. Die Werte auf dem Kreisumfang werden für $|\underline{r}_a| = 1$ erreicht, d. h. bei einem total reflektierenden Verbraucher (Reaktanzzweipol). Die Punkte kleinsten Betrages (linker Schnittpunkt des Kreises mit der reellen Achse bei $1 - |\underline{r}_i|$) und größten Betrages (rechter Schnittpunkt bei $1 + |\underline{r}_i|$) entsprechen gemäß Gl. (12) der ablaufenden Welle \underline{b} mit größtem bzw. kleinstem Betrag. Bei einem passiven Abschluß gibt es also für den Betrag $|\underline{b}|$ der vom Generator ablaufenden Welle eine obere und untere Grenze:

$$|\underline{b}|_{\max} = \frac{|\underline{b}_q|}{1 - |\underline{r}_i|}, \quad |\underline{b}|_{\min} = \frac{|\underline{b}_q|}{1 + |\underline{r}_i|}. \quad (13)$$

³ Wegen $|\underline{r}_a| \leq 1$ ist $|\underline{r}_a \underline{r}_i| \leq |\underline{r}_i|$, $\underline{r}_a \underline{r}_i$ und auch $-\underline{r}_a \underline{r}_i$ liegen also innerhalb eines Kreises um den Nullpunkt mit dem Radius $|\underline{r}_i|$.

Diese Extremwerte werden nur bei Totalreflexion ($|\underline{r}_a| = 1$) erreicht. Aus Gl. (12) ergibt sich, daß das Maximum von $|\underline{b}|$ dann auftritt, wenn $\underline{r}_i \underline{r}_a$ positiv reell ist, d. h.

$$\arccos \underline{r}_i + \arccos \underline{r}_a = 0 \quad (14)$$

ist. Entsprechend gilt für das Minimum von $|\underline{b}|$, daß $\underline{r}_i \underline{r}_a$ negativ reell ist, d. h.

$$\arccos \underline{r}_i + \arccos \underline{r}_a = \pi \quad (15)$$

ist.

Aus Gl. (13) läßt sich ein einfaches Verfahren herleiten, wie man bei einem Generator mit Koaxial- oder Hohlleiterausgang den Betrag des inneren Reflexionsfaktors \underline{r}_i mißt. Man schließt an den Generator über einen Richtkoppler einen veränderbaren Reaktanzzweipol an, z. B. in Form einer Kurzschlußleitung einstellbarer Länge (Bild 7). Mit Hilfe des Richtkopplers wird der Betrag $|\underline{b}|$ der hin-

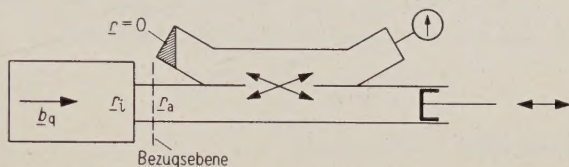


Bild 7. Anordnung zur Messung von $|\underline{r}_i|$ bei einem Generator mit Hohlleiterausgang.

laufenden Welle gemessen, der in Abhängigkeit von der Stellung des Kurzschlußschiebers zwischen den Werten $|\underline{b}|_{\min}$ und $|\underline{b}|_{\max}$ hin- und herpendelt. Mit Gl. (13) ergibt sich aus diesen Extremwerten von $|\underline{b}|$ der Betrag $|\underline{r}_i|$ des inneren Reflexionsfaktors zu

$$|\underline{r}_i| = \frac{|\underline{b}|_{\max} - |\underline{b}|_{\min}}{|\underline{b}|_{\max} + |\underline{b}|_{\min}}. \quad (16)$$

Mißt man außerdem in der Stellung des Kurzschlußschiebers, bei der $|\underline{b}|_{\max}$ auftritt, den Winkel des Eingangsreflexionsfaktors \underline{r}_a in der Bezugsebene, z. B. mit Hilfe einer Meßleitung, so läßt sich daraus mit Gl. (14) auch der Winkel des inneren Reflexionsfaktors \underline{r}_i bestimmen.

Wir berechnen nun die Wirkleistung P_p , die der Generator an den Verbraucher abgibt. Mit den Gl. (6), (11) und (12) ergibt sie sich zu

$$P_p = \frac{1}{2} |\underline{b}|^2 (1 - |\underline{r}_a|^2), \quad (17)$$

$$P_p = \frac{1}{2} \frac{|\underline{b}_q|^2 (1 - |\underline{r}_a|^2)}{|1 - \underline{r}_i \underline{r}_a|^2}.$$

Die Wirkleistung P_p ist bei gegebenem Generator ($\underline{b}_q, \underline{r}_i$) vom Verbraucher (\underline{r}_a) abhängig. Im Anhang wird gezeigt, daß in der \underline{r}_a -Ebene die Kurven konstanter Wirkleistung Kreise sind. Wie eine Extremwertbetrachtung zeigt, tritt das Maximum von P_p auf, wenn $\underline{r}_a = \underline{r}_i^*$ ist⁴, d. h. die Reflexionsfaktoren

⁴ Man sieht zunächst, daß in bezug auf den Winkel von \underline{r}_a ein Maximum auftritt, wenn $\arccos \underline{r}_i = -\arccos \underline{r}_a$ ist. Der verbleibende Ausdruck

$$\frac{1}{2} |\underline{b}_q|^2 \frac{1 - |\underline{r}_a|^2}{(1 - |\underline{r}_i| |\underline{r}_a|)^2}$$

ist nur noch von $|\underline{r}_a|$ abhängig und nimmt bei $|\underline{r}_a| = |\underline{r}_i|$ sein Maximum an.

von Generator und Verbraucher konjugiert komplex sind. Dieser Fall ist die bekannte Wirkleistungsanpassung einer Zweipolquelle, bei der Generatorinnenwiderstand Z_i und Verbraucherwiderstand Z_a konjugiert komplex sind. Die maximale Wirkleistung $P_{p\max}$ ergibt sich dabei zu

$$P_{p\max} = \frac{1}{2} \frac{|b_q|^2}{1 - |r_i|^2}. \quad (18)$$

Den Betrag der bei Leistungsanpassung vom Generator ablaufenden Welle b_A erhalten wir aus Gl. (12) mit $r_a = r_i^*$ zu

$$|b_A| = \frac{|b_q|}{1 - |r_i|^2}. \quad (19)$$

Der Betrag der bei *Wirkleistungsanpassung* ablaufenden Welle b_A ist also stets größer als der Betrag der bei *Wellenanpassung* ($r_a = 0$) ablaufenden Urwelle b_q . Wir bemerken schließlich noch, daß mit Gl. (19) die Maximalleistung auch in der Form

$$P_{p\max} = \frac{1}{2} |b_q| |b_A| \quad (20)$$

geschrieben werden kann.

3. Transformation der Ersatzwellenquelle über Leitungen und Vierpole

Wenn an die Klemmen eines Generators (gekennzeichnet durch die Größen b_q und r_i) eine homogene Leitung mit der Länge l und dem bezogenen Phasenmaß β angeschlossen wird (Bild 8), deren Wellenwiderstand außerdem gleich dem Normierungswiderstand Z_L ist, so läßt sich für das Leitungsende sehr einfach eine neue Ersatzwellenquelle angeben.

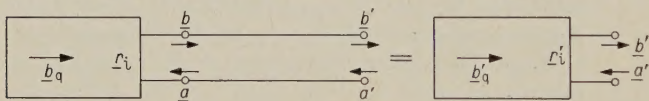


Bild 8. Transformation der Ersatzwellenquelle über eine Leitung.

Bezeichnen wir mit b' gemäß Bild 8 die vom Leitungsende nach rechts ablaufende Welle und mit a' die zulaufende Welle, so gilt

$$b' = b e^{-j\beta l}, \quad a' = a e^{+j\beta l}. \quad (21)$$

Für den linken Generator gilt

$$b = b_q + r_i a. \quad (22)$$

Mit Gl. (21) erhält man für die ausgangsseitigen Wellen die Beziehung

$$\begin{aligned} b' e^{j\beta l} &= b_q + r_i a' e^{-j\beta l}, \\ b' &= \underbrace{b_q e^{-j\beta l}}_{b'_q} + \underbrace{r_i e^{-2j\beta l} a'}_{r'_i}. \end{aligned} \quad (23)$$

Die Urwelle b'_q des Ersatzgenerators für das Leitungsende eilt also um den Winkel βl gegenüber der Urwelle b_q des ursprünglichen Generators nach und hat den gleichen Betrag. Der innere Reflexions-

faktor r'_i geht aus r_i durch eine Drehung um den Winkel $2\beta l$ im Uhrzeigersinn hervor.

Ebenso läßt sich für die Ausgangsklemmen eines allgemeinen linearen Vierpols eine Ersatzwellenquelle angeben, wenn man seine Eingangsklemmen mit einem Generator (b_q, r_i) gemäß Bild 9 verbindet.

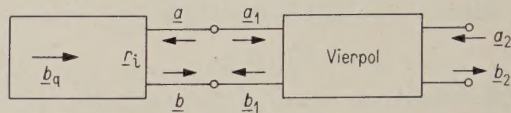


Bild 9. Zusammenschaltung eines Generators und eines Vierpols.

Der Vierpol sei durch seine Streumatrix beschrieben:

$$b_1 = S_{11} a_1 + S_{12} a_2, \quad (24a)$$

$$b_2 = S_{21} a_1 + S_{22} a_2. \quad (24b)$$

Für den linken Generator gilt

$$b = b_q + r_i a$$

oder mit $a = b_1$ und $b = a_1$:

$$a_1 = b_q + r_i b_1. \quad (25)$$

Durch Einsetzen von Gl. (25) in Gl. (24a) erhält man

$$a_1 = \frac{S_{12} a_2 + (b_q/r_i)}{(1/r_i) - S_{11}}$$

und daraus in Verbindung mit Gl. (24b) folgende Beziehung für die Wellengrößen an den Ausgangsklemmen:

$$b_2 = \underbrace{\frac{S_{21}}{1 - r_i S_{11}} b_q}_{b'_q} + \underbrace{\left(\frac{S_{22} + r_i S_{12} S_{21}}{1 - r_i S_{11}} \right) a_2}_{r'_i}. \quad (26)$$

Mit dieser Gleichung lassen sich die Daten der ausgangsseitigen Ersatzwellenquelle (Urwelle b'_q und innerer Reflexionsfaktor r'_i) berechnen. Für den Sonderfall der homogenen Leitung mit $S_{11} = S_{22} = 0$ und $S_{12} = S_{21} = e^{-j\beta l}$ geht Gl. (26) in Gl. (23) über.

Eine besondere Bedeutung hat die Transformation eines Generators über einen *verlustfreien* Vierpol. In diesem Sonderfall sind die maximalen Wirkleistungen, die man aus dem ursprünglichen Generator (b_q, r_i) und dem ausgangsseitigen Ersatzgenerator (b'_q, r'_i) entziehen kann, einander gleich. Nach Gl. (18) gilt also für verlustfreie Vierpole

$$\frac{|b'_q|^2}{1 - |r'_i|^2} = \frac{|b_q|^2}{1 - |r_i|^2}. \quad (27)$$

Häufig liegt die Aufgabe vor, einen allgemeinen Generator (b_q, r_i) durch einen verlustfreien Vierpol in eine reine Urwellenquelle ($r'_i = 0$) zu transformieren. Aus Gl. (27) entnimmt man, daß in diesem Fall

$$|b'_q|^2 = \frac{|b_q|^2}{1 - |r_i|^2} \quad (28)$$

ist, d. h. die ausgangsseitige Urwelle ist größer als die eingangsseitige. Ebenso sieht man, daß diese Ur-

welle größer ist als die bei allen anderen verlustfreien Transformationen gemäß Gl. (27) erzielbare Urwelle. Wir können aus dieser Überlegung auch herleiten, unter welcher Bedingung die Transformation in eine reine Urwellenquelle erfolgt. Die Größe

$$|b'_q|^2 = \frac{|S_{21}|^2}{|1 - r_i S_{11}|^2} |b_q|^2$$

(nach Gl. (26)) muß nämlich bei Variation der Größen S_{11} und S_{21} möglichst groß werden. Beachtet man noch, daß bei verlustfreien Vierpolen die Beziehung

$$|S_{11}|^2 + |S_{21}|^2 = 1$$

(siehe z. B. [2]) gilt, muß also der Ausdruck

$$\frac{1 - |S_{11}|^2}{|1 - r_i S_{11}|^2}$$

möglichst groß werden. Ein analoger Ausdruck kam bereits in Gl. (17) vor. Das Maximum tritt also entsprechend bei

$$S_{11} = r_i^* \quad (29)$$

auf. Unter dieser Bedingung wird also ein Generator mit dem inneren Reflexionsfaktor r_i durch einen verlustfreien Vierpol in eine reine Urwellenquelle transformiert.

4. Technische Generatoren

Die Generatoren der Nachrichtentechnik sind zum großen Teil keine *linearen* Zweipolquellen. Beispiele dafür sind der rückgekoppelte Röhrenoszillator, das Reflexklystron und das selten gebrauchte Dynatron. Hier beruht sogar der Vorgang der Schwingungserzeugung wesentlich auf einer nicht-linearen Kennlinie. Für Zweipolquellen, deren Klemmen unmittelbar mit einem solchen Oszillator verbunden sind, läßt sich keine Ersatzwellenquelle angeben. Hier fehlt nicht nur der lineare Zusammenhang zwischen Klemmenspannung und -strom, vielmehr ist auch die Generatorfrequenz belastungsabhängig. Diese Generatoren können nur dadurch vollständig beschrieben werden, daß man für jede Belastung die Frequenz und die abgegebene Leistung angibt (in der Mikrowellentechnik z. B. in Form eines Rieke-Diagramms).

Die *lineare* Zweipolquelle läßt sich jedoch technisch mit guter Genauigkeit dadurch realisieren, daß man die Zweipolklemmen vom eigentlichen Oszillator „isoliert“. In der Niederfrequenztechnik geschieht das in der Regel dadurch, daß man zwischen dem Oszillator und die Klemmen eine rückwirkungs-freie Verstärkerstufe schaltet. In der Mikrowellentechnik wird die Isolation meistens durch einen Richtungsleiter vorgenommen, der im Inneren einen Werkstoff mit nichtreziprokem Verhalten (z. B. Ferrit) enthält und nur Energie vom Oszillator zu den Klemmen, aber nicht in umgekehrter Richtung überträgt. In den Frequenzbereichen, in denen die Röhrenverstärker nicht mehr rückwirkungsfrei arbeiten und Ferrit-Richtungsleiter noch nicht zur Verfügung stehen, muß man zur Ent-

kopplung von Oszillator und Zweipolklemmen einen reziproken Vierpol hoher Dämpfung einfügen. Die Isolation muß dann allerdings mit einem Leistungsverlust erkauft werden.

Anhang

Mit Gl. (17) und Gl. (18) ergibt sich für die auf den Maximalwert $P_{p \max}$ bezogene Wirkleistung P_p

$$\frac{P_p}{P_{p \max}} = \frac{(1 - |r_a|^2)(1 - |r_i|^2)}{|1 - r_a r_i|^2} \quad (A.1)$$

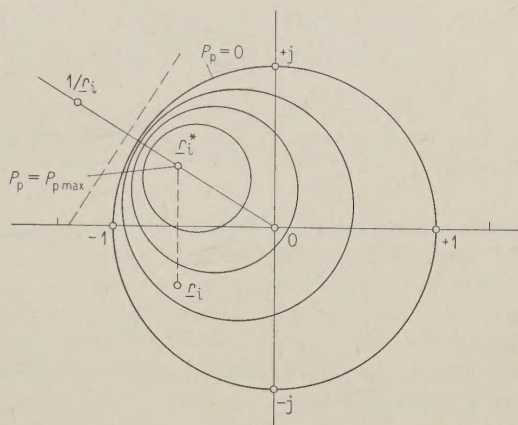


Bild 10. Kreise konstanter Wirkleistung in der r_a -Ebene (r_a Reflexionsfaktor des Verbrauchers).

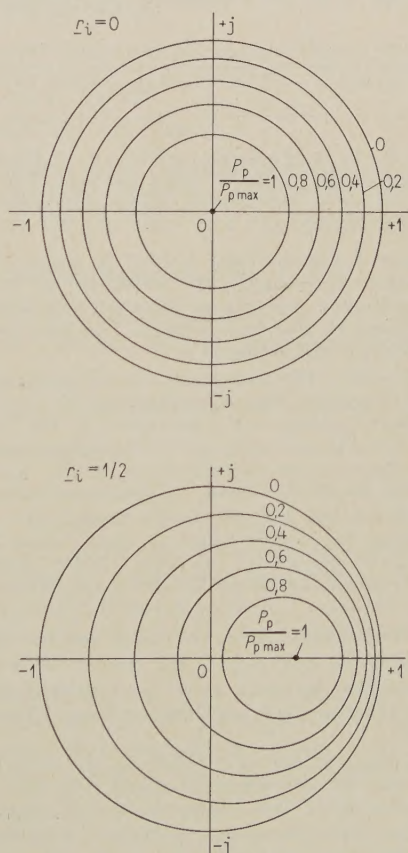


Bild 11. Kreise konstanter Wirkleistung für $r_i = 0$ und $r_i = 1/2$.

Dann gilt

$$1 - \frac{P_p}{P_{p \max}} = \frac{|1 - r_a r_i|^2 - (1 - |r_a|^2)(1 - |r_i|^2)}{|1 - r_a r_i|^2} \quad (\text{A. 2})$$

Der Zähler wird in folgender Weise umgeformt:

$$\begin{aligned} |1 - r_a r_i|^2 - (1 - |r_a|^2)(1 - |r_i|^2) &= \\ &= (1 - r_a r_i)(1 - r_a^* r_i^*) - 1 + r_a r_a^* + \\ &\quad + r_i r_i^* - r_a r_i r_a^* r_i^* = \\ &= r_a r_a^* + r_i r_i^* - r_a r_i - r_a^* r_i^* = \\ &= (r_a - r_i^*)(r_a^* - r_i) = \\ &= |r_a - r_i^*|^2. \end{aligned}$$

Aus dem Nenner von Gl. (A. 2) wird $|r_i|^2$ ausgeklammert:

$$|1 - r_a r_i|^2 = |r_i|^2 |r_a - (1/r_i)|^2.$$

Damit ergibt sich aus Gl. (A. 2)

$$1 - \frac{P_p}{P_{p \max}} = \frac{1}{|r_i|^2} \frac{|r_a - r_i^*|^2}{|r_a - (1/r_i)|^2} \quad (\text{A. 3})$$

Die Kurven konstanter Wirkleistung in der r_a -Ebene erhält man, wenn

$$\frac{|r_a - r_i^*|}{|r_a - (1/r_i)|}$$

konstant ist, d. h. das Verhältnis der Abstände eines

Punktes r_a von den festen Punkten r_i^* und $1/r_i$ konstant ist. Diese Kurven sind sogenannte Apollonius-Kreise (Bild 10). Ihre Mittelpunkt liegen alle auf einer Geraden durch den Nullpunkt und die Punkte r_i^* und $1/r_i$. Diese Kreise ergeben sich aus den bekannten WEISSFLOCHSchen Kreisen konstanter Wirkleistung in der Widerstandsebene (Z_a -Ebene)

[7] durch die Abbildung $r_a = \frac{Z_a - Z_L}{Z_a + Z_L}$. Bild 11 zeigt zwei typische Kreisscharen konstanter Wirkleistung, nämlich für $r_i = 0$ und $r_i = 1/2$.

An dieser Stelle möchte ich Herrn Prof. Dr.-Ing. H. DÖRING und meinen Kollegen, Herrn H. BRAND und Herrn Th. A. ABELE, für viele Anregungen danken.

Schrifttum

- [1] WALLOT, J., Einführung in die Theorie der Schwachstromtechnik, 5. Auflage. Springer-Verlag, Berlin 1948, §§ 15–24, § 107.
- [2] SCHUON, E. und WOLF, H., Die Darstellung von Mehrpolen durch die Streumatrix. Nachrichtentech. Z. **12** [1959], 361–366, 408–415.
- [3] HELMHOLTZ, H., Über einige Gesetze der Verteilung elektrischer Ströme in körperlichen Leitern mit Anwendung auf die tierisch-elektrischen Versuche. Poggendorffs Ann. **89** [1853], 211–230.
- [4] MAYER, H. F., Über das Ersatzschema der Verstärkerröhre. Telegr. u. Fernspr. Tech. **15** [1926], 335–337.
- [5] FELDTKELLER, R., Einführung in die Vierpoltheorie der elektrischen Nachrichtentechnik, 7. Auflage. S. Hirzel Verlag, Stuttgart 1959, S. 4–12.
- [6] Normblatt DIN 1323, Elektrische Spannung, Potential, Zweipolquelle, elektromotorische Kraft.
- [7] WEISSFLOCH, A., Die Wirkleistungsabgabe eines Generators mit konstantem Innenwiderstand und konstanter EMK an beliebige komplexe Außenimpedanzen. Hochfrequenztech. u. Elektroakustik **60** [1942], 10–11.

BUCHBESPRECHUNGEN

R. Rost, Kristallogudentechnik; 2. Ergänzungsband, 2. Auflage. Verlag W. Ernst und Sohn, Berlin 1960, VIII, 166 Seiten, 191 Bilder, 15 cm × 21,5 cm, Ganzleinen DM 28,—.

In diesem Ergänzungsband werden in erster Linie die Schalttransistoren behandelt. Die verschiedenen Typen und ihre Eigenschaften werden in übersichtlicher Weise dargestellt, ebenso wie ihre Anwendungen. Dies geht aus folgenden Kapitelüberschriften hervor:

1. Der Begriff, 2. Die Analyse des Schalt-Transistors, 3. Die Triggerschaltungen, 4. Die Schalt-Triode mit doppelter Basis, 5. Der Schalt-Transistor mit Zusatzspitze, 6. Ein Zähl- und Steuergerät mit Transistoren und Kaltkathodenröhren, 7. Die Transistor-Bauteile des Stellenrechners, 8. Die ersten transistorisierten Rechenmaschinen, 9. Die Registrierung einer Stabverrückung mit Transistor und Photo-Diode, 10. Die Zähl-schaltung mit Sichtanzeige, 11. Der Analog-Rechner, 12. Die Schalttriode als HF-Oszillator, 13. Der Schalt-Transistor in Spannungswandlern, 14. Die elektronischen Schaltgeräte.

Der Text wird durch eine große Anzahl von wertvollen Abbildungen ergänzt.

Der zweite Hauptabschnitt des Buches behandelt eine Reihe von Gegenständen, welche aus folgenden Kapitelüberschriften hervorgeht:

15. Die Eigenschaften des Drifttransistors bei UKW, 16. Der Spacistor, 17. Der diffusionslegierte Transistor, 18. Der Doppeltransistor, 19. Ein kleines Transistor-Prüfgerät, 20. Der passive Transistorempfänger, 21. Der Sekundärsender, 22. Die Transistorkennwerte h_{11} , h_{12} , h_{21} und h_{22} .

Der dritte Hauptabschnitt enthält die Beschreibung einer Reihe von Firmenerzeugnissen. Die Kapitel sind:

23. Schalt-Dioden, 24. Schalt-Transistoren, 25. Photo-Kristallogen, 26. Leistungs-Transistoren, 27. Hochfrequenz-Transistoren, 28. Ge-Niederfrequenz-Transistoren, 29. Transistorsätze für Kleinstradios, 30. Miniaturtransistoren, 31. Siliziumtransistoren, 32. Firmenverzeichnis, 33. Herstellerfirmen.

Das Buch ist eine wertvolle Ergänzung der Literatur über Transistoren und für jeden Benutzer von Transistoren fast unentbehrlich.

M. STRUTT

K. M. Koch und R. Reinbach, Einführung in die Physik der Leiterwerkstoffe. Verlag F. Deuticke, Wien 1960, VIII, 255 Seiten, 149 Bilder, 29 Tabellen, 17 cm × 24,5 cm, Ganzleinen DM 37,—.

Das Buch bringt in einem physikalischen Teil die theoretischen Grundlagen der Elektrizitätsleitung von der klassischen bis zur wellenmechanischen Elektronentheorie, die Leiteigenschaften von Metallen und Legierungen, die Supraleitung, die thermoelektrischen Effekte und den Einfluß von Magnetfeldern. Die Halbleiter werden bei der Erläuterung des Energiebänder-Modells gestreift. Auf etwa 80 Seiten werden die für die Elektrotechnik wichtigen Leiterwerkstoffe mit ihren elektrischen und mechanischen Eigenschaften und ihrem Korrosionsverhalten behandelt. Hier findet man viele Daten über gute Leiter und über Legierungen für Widerstände und Thermoelemente, die durch metallkundliche Hinweise verständlich gemacht werden. Damit bietet diese Einführung nicht nur eine sehr gute Informationsmöglichkeit über metallische Leitung und Leiter für Physiker und Ingenieure — in der Gegenüberstellung von Theorie und Praxis weist sie auch auf die Probleme hin, deren Verfolgung im Laboratorium und am Schreibtisch wichtig und reizvoll erscheint.

K. SIXTUS

Aktive Hochfrequenzspektrometer für ionosphärische Echolotung*

II. Selektion der Echozeichen aus dem Störpegel

VON KARL RAWER

Ionosphären-Institut Breisach im Fernmeldetechnischen Zentralamt der Deutschen Bundespost

(A.E.Ü. 14 [1960], 373–379; eingegangen am 24. Juli 1960)

DK 551.508.96

Die a-priori-Kenntnis der Echoimpulse wird bisher nur zum geringen Teil genutzt. Mögliche Anwendungen werden für vier Kriterien diskutiert, nämlich: Periodizität, Form des Signals, hochfrequente Phase, Einfallsrichtung. Durch bessere Nutzung solcher Unterscheidungs-Kriterien sollten Verbesserungen zu erzielen sein.

The a priori knowledge concerning the echo pulses has only partially been used in the past. Possible applications are discussed for four criteria, viz.: Periodicity, Signal shape, rf phase, Direction of arrival. With a better use of such distinctive criteria it should be possible to obtain improvements.

Einleitung

Beim Empfang von Echoimpulsen in einem breiten Frequenzbereich, wie er z. B. bei der ionosphärischen Echolotung in Betracht kommt, treten durch die mannigfaltige Belegung der überstrichenen Bänder an vielen Stellen starke Störsignale auf. Die praktische Begrenzung der Wiedergabe liegt bei der Technik der Impulslotung mit kontinuierlich veränderter Frequenz deshalb nicht bei den natürlichen Störgeräuschen, sondern bei den wesentlich stärkeren, künstlichen Sendungen.

Die Informationstheorie lehrt, daß zur Unterscheidung zwischen einem echten Signal und einem Störzeichen die a-priori-Kenntnis des echten Signals von ausschlaggebender Bedeutung ist. Je mehr a priori über die echten Signale bekannt ist, desto besser kann man sie von Störzeichen unterscheiden, und jedes Kriterium, das spezifisch auf die echten Signale zutrifft, kann zur Unterscheidung herangezogen werden.

Einige der im folgenden genannten Methoden sind seit langem bekannt und werden üblicherweise beim Impulsempfang angewandt, die meisten jedoch sind bisher höchstens in speziellen Anwendungen benutzt worden. Ziel der folgenden Diskussion ist es, festzustellen, wie weit genauere Spezifizierung zur praktischen Empfangsverbesserung ausgenutzt werden kann.

1. Mögliche Kriterien

Stellen wir zunächst die wichtigsten Eigenschaften zusammen, die wir a priori den von der Ionosphäre kommenden Echos zuschreiben können. Sie ergeben sich aus unserer genauen Kenntnis der Sendeimpulse einerseits und aus der Lage der reflektierenden ionosphärischen Schichten andererseits. Im einzelnen wollen wir die folgenden Kriterien diskutieren:

1.1. Periodizität

Die Wiederholungsperiode τ der Sendeimpulse ist bekannt und sollte sich in den Echos wiederholen,

sofern die Reflexionshöhe sich innerhalb der Wiederholungsperiode nicht wesentlich verändert. Während also die Wiederholungsfrequenz der Echos als bekannt vorausgesetzt werden kann, ist deren niederfrequente Phase, die der Reflexionshöhe entspricht, nahezu unbekannt. Für Ionosphärenlotungen können wir nur feststellen, daß die interessierenden Echos im allgemeinen Laufzeiten von 0,6 bis 4 ms (scheinbare Höhen von 90 bis 600 km) haben können.

1.2. Form des Signals

Grundsätzlich sind Aussagen möglich über a) die Dauer des Signals, b) seine Form, c) seine Amplitude.

1.3. Hochfrequente Phase

Abgesehen von einem Phasensprung, der der Phasenlaufzeit entspricht, ist sie für jedes Echo kohärent mit der Phase des ausgesandten Zeichens.

1.4. Einfallsrichtung der Echos

Die interessierenden Echos von der Ionosphäre kommen praktisch alle aus vertikaler Richtung.

2. Anwendung der Kriterien

2.1. Periodizität

Die Periodizität der Echos wird praktisch bei allen Ionosphärenregistrierungen ausgenutzt. Statt einfach den Zeitablauf der Ausgangsspannung des Empfängers zu registrieren und daraus die Echozeichen einzeln herauszusuchen, synchronisiert man die Ablenkspannung der Oszillographenröhre mit den Sendeimpulsen und registriert photographisch mit Überschiebung. Man erreicht dadurch einerseits, daß während der Überschiebungszeit T alle anfallenden Echozeichen, der Zahl nach T/τ , dieselbe Stelle des Films schwärzen, während die zufällig verteilten Störsignale sich nicht in gleicher Weise addieren, sondern bei genügend langer Über-

* Die Arbeit wurde durch eine Forschungsbeihilfe des Bundesministeriums für Wirtschaft, Nr. A 190, ermöglicht.

schreibungszeit einen gleichmäßigen Schwärzungsuntergrund liefern. Andererseits erreicht man dadurch eine bessere Lesbarkeit im Spektrum, weil die Spur der Echozeichen nun stetig verläuft und deshalb leicht über stark gestörte Frequenzen hinweg interpoliert werden kann. Der Gewinn, der durch Impulswiederholung erzielbar ist, und die Folgerungen, die sich daraus für die Dimensionierung aktiver Hochfrequenzspektrometer ergeben, werden im Abschnitt 3.1 genauer besprochen.

Einen interessanten Versuch, die niederfrequente Phase trotz der unbekannten Lage der Echoimpulse als Kriterium heranzuziehen, hat BIBL [1] unternommen. Er führt im interessierenden Höhenbereich eine Farbkennzeichnung ein und registriert auf Farbfilm. Je nach der Laufzeit erscheint das Echo in anderer Farbe als der farblich gemischte Störpegel. Mit variabler Echohöhe verändert sich zwar die Farbe des Echos, jedoch in stetiger Weise.

2.2. Form des Signals

2.2.1. Dauer des Signals

Die Dauer eines Echsignals, das von einem klaren Reflexionsniveau kommt, ist durch die Dauer des ausgesandten Zeichens und durch die Wiedergabeeigenschaften, besonders die Bandbreite des Empfängers [2], bestimmt. Grundsätzlich könnte sie zur Auslese der richtigen Echozeichen benutzt werden, wenn das Ziel eindeutig spiegelte. Das ist zwar bei vielen Echos aus der Ionosphäre der Fall, jedoch nicht bei allen. Speziell durch die magnetische Aufspaltung ergeben sich im Spektrum immer breitere Bereiche, in denen die Echodauer scheinbar über das oben angegebene Maß hinaus erhöht ist. Deshalb besteht für den speziellen Fall der Ionosphärenlotungen keine Möglichkeit, die Signaldauer als Kriterium auszunutzen.

2.2.2. Form des Signals

Die Form eines Impulssignals nach Passieren des Empfängers ist durch einen mehr oder minder steilen Anstieg zu Beginn gekennzeichnet [2]. Dieser Anstieg wird häufig als Kriterium verwendet. Beschneidet man nämlich den Niederfrequenzverstärker des Empfängers nach niederen Frequenzen zu, so werden rasche Veränderungen relativ stärker wiedergegeben als langsame.

Man erreicht das im allgemeinen durch ein differenzierendes RC -Glied. Dadurch werden Dauerstrichstörungen stark verringert, so daß man eine deutlichere Wiedergabe der Echospur erhält. Die Methode des Differenzierens hat jedoch vom Standpunkt der Informationsbetrachtung ganz erhebliche Nachteile. Einerseits geht beim Differenzieren die Kenntnis der absoluten Amplitude des Echos verloren; höchstens die relative Amplitude im Vergleich zur jeweiligen Störampplitude wird noch wiedergegeben. Außerdem wird durch Differenzieren ein Laufzeitfehler eingeführt, der seinerseits von der Störampplitude abhängt. Deshalb erzeugen Störer eine scheinbare Schwankung der Laufzeit und die Messung wird so erheblich ungenauer (siehe

Bild 1). Schließlich entsteht beim Differenzieren hinter starken Echsignalen ein totes Zeitintervall, innerhalb dessen schwache Signale unterdrückt werden. So geht durch Differenzieren wertvolle Information verloren.

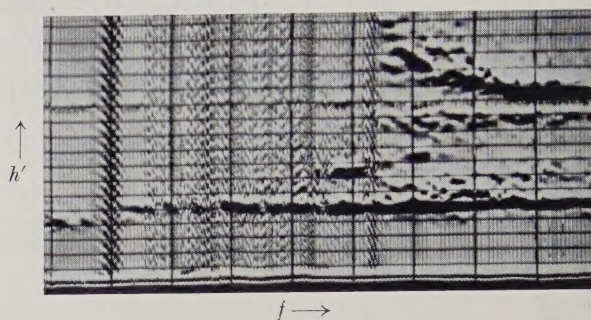


Bild 1. Echoregistrierung mit Differenzierung; Laufzeitfehler durch Fremdstörer [dunkle Zeichen auf hellem Grund] (Ionogramm = Echoregistrierung mit ansteigender Frequenz) im Mittelwellenbereich, National Bureau of Standards, Boulder, Col., U.S.A.).

Die angegebenen Nachteile sind nicht unvermeidlich. Eine Möglichkeit, sie wenigstens teilweise zu vermeiden, wird im Abschnitt 3.2 angegeben.

2.2.3. Amplitude des Signals

Während bei manchen technischen Anwendungen des Radarverfahrens der Bereich der interessierenden Amplituden von vornherein angegeben werden kann, ist man bei anderen Anwendungen gerade daran interessiert, auch schwächste Echsignale noch zu empfangen. Für die Echolotung der Ionosphäre könnte man die Meinung vertreten, allzu schwache Signale seien ohne Interesse. Ziel der Echolotungen ist es ja, den Ionosphärenzustand über dem Beobachtungsort zu beschreiben, ohne auf Kleinigkeiten einzugehen [3]. Wohl aus diesem Grunde benutzen viele Echolotungsgeräte eine Begrenzung der Empfangssignale und schneiden alle Signale unter einer gewissen Grenze ab. Damit erhält man „schön aussehende“ Registrierungen, bei denen die Echozeichen (allerdings auch die starken Störer) als klare Spuren auf glattem Grund erscheinen. Bild 2 zeigt im Vergleich mit dem nahezu gleichzeitig aufgenommenen Bild 3, bei dem keine Begrenzung benutzt wurde, daß auch hier wertvolle Information verloren geht. Zwar kann man die Amplitude der Echozeichen, von der unteren Begrenzung abgesehen, noch stetig wiedergeben, wie wir es im Bild 2 getan haben, aber es gehen, je nach Einstellung der Begrenzung, Teile der Echospur verloren, die von erheblichem Interesse wären. Durch die auf niederen Frequenzen besonders wirksame Absorption in den tieferen Schichten der Ionosphäre sinkt bei Tag die Amplitude mit abnehmender Frequenz; deshalb wird die untere Frequenzgrenze des Echobereichs durch eine Begrenzung merklich heraufgesetzt. Außerdem zeichnen sich bei Wiedergabe mit Begrenzung kurzzeitige Fadings als Lücken in der Spur ab und erhalten dadurch ein

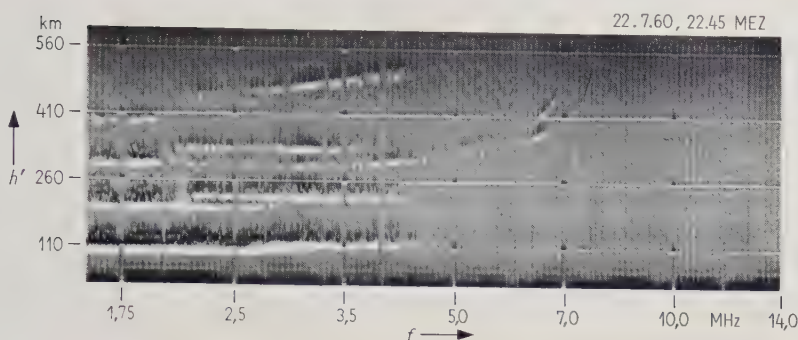


Bild 2. Echoregistrierung mit Amplitudenselektion, stark differenziert [helle Zeichen auf dunklem Grund] (Ionogramm im Kurzwellenbereich, Ionosphären-Institut Breisach).

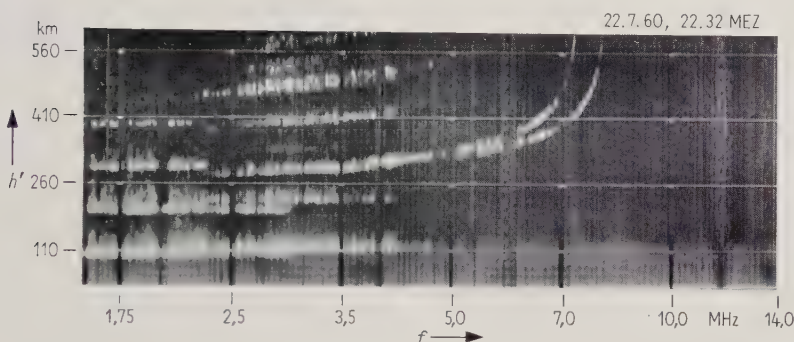


Bild 3. Echoregistrierung ohne Amplitudenselektion, nur schwach — d. h. mit großer Zeitkonstante — differenziert [helle Zeichen auf dunklen Grund] (Ionogramm im Kurzwellenbereich, Ionosphären-Institut Breisach).

übertriebenes Gewicht. Aus diesen Gründen muß die Amplitudenbegrenzung für ionosphärische Registrierungen als ungeeignet verworfen werden.

Eine ähnliche Wirkung, ohne die schädlichen Nebenwirkungen, hat die gelegentlich angewandte Technik der zusätzlichen Helligkeitssteuerung, die sich allerdings nur bei Amplitudenregistrierungen der Echos verwerten läßt. Man gibt grundsätzlich die Ausgangsspannung des Empfängers einerseits auf die Ablenkplatten der Oszillographenröhre, andererseits auf das Helligkeitsgitter. Starke Echozeichen werden der Helligkeit nach kräftiger wiedergegeben und erscheinen deshalb deutlicher. Leider erlaubt die Methode keine unmittelbare Anwendung auf Hochfrequenzspektren, weil dort keine seitliche Auslenkung zur Amplitudenwiedergabe mehr zur Verfügung steht. (Vergleiche hierzu den Vorschlag von BIBL [1], die Farbe als zusätzlichen Parameter heranzuziehen.)

2.3. Hochfrequente Phase

Bei den üblichen Empfängern wird die Kenntnis der hochfrequenten Phase bis zur letzten Zwischenfrequenzstufe mitgeführt, geht jedoch dann in der Gleichrichtung verloren. Nur in speziellen Anwendungen auf fester Frequenz hat man bisher versucht, die hochfrequente Phase in die Registrierung hinüber zu retten. Das Verfahren ist zunächst von RANZI [4] angegeben und später von FINDLAY [5] sowie McNICOL [6] verbessert worden. Man startet

gleichzeitig mit dem Sendepuls einen Oszillator, dessen Frequenz knapp neben der Sendefrequenz liegt; die Schwingung des Oszillators wird erst kurz vor dem nächsten Impuls durch einen Sperrimpuls gestoppt, so daß sie dann beim kommenden Impuls, unabhängig von der Endphasenlage, wieder beginnen kann. Die Schwingung dieses Oszillators wird, passend abgeschwächt, zusätzlich in den Empfängereingang eingekoppelt. Durch Interferenz entsteht eine Modulation des direkten Zeichens mit der Differenzfrequenz. Bei der üblichen Wiedergabe der Echozeichen erscheinen in diesen Interferenzstreifen, die sich mit der Phasenlaufzeit ändern. Man erhält so Registrierungen mit schrägläufigen Interferenzstreifen. Auch wenn man nicht an Details der Phasenlaufzeit interessiert ist, ermöglicht die Streifung eine weit bessere Unterscheidung von Echo und Störpegel; außerdem lassen sich schwache Echozeichen verschiedener Herkunft mit dieser Methode trennen, während sie bei der üblichen Helligkeitsregistrierung ineinander übergehen. Ein Beispiel zeigt Bild 4 [6]. Man erkennt im oberen Teil sehr schwache Vielfachechos, die andernfalls wohl im Störpegel verschwinden würden. Im unteren Teil sieht man zwei verschiedene Es-Echos, die durch die verschiedene Streifung sich deutlich unterscheiden, während sie andernfalls ineinander übergehen würden. Der Informationsgewinn durch Phasenunterscheidung ist also deutlich.

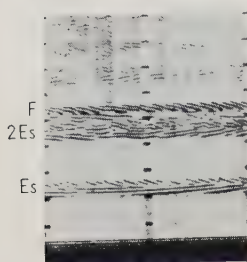


Bild 4. Echoregistrierung mit Phasenanzeige [dunkle Zeichen auf hellem Grund] (feste Frequenz, University of Brisbane, Australien).

Bei der Registrierung mit Interferenzstreifen geht allerdings etwas an Information über die Gruppenlaufzeit verloren, weil ja die Unterkante des Echozeichens nun ausgefranst und damit undeutlicher wiedergegeben wird. McNICOL [6] hat den Hilfsoszillator an die Zwischenfrequenz und nicht an die ausgesandte Frequenz „angebunden“, was in der technischen Anwendung eine erhebliche Vereinfachung bedeutet. Möglichkeiten der Anwendung für Hochfrequenzspektrometer werden im Abschnitt 3.3 diskutiert.

2.4 Einfallrichtung der Echos

Echos von der Ionosphäre fallen überwiegend aus vertikaler Richtung in einem Winkelbereich von $\pm 5^\circ$ ein. Für die Erfordernisse einer Echolotungsstation ist es gar nicht erwünscht, Echos aus schräger Richtung zu erhalten [1]. Deshalb wäre eine scharfe Bündelung in vertikaler Richtung sender- und empfängerseitig zweckmäßig und könnte eine erhebliche Verbesserung des Abstandes vom Störpegel bringen. Leider sind aber im Frequenzbereich der ionosphärischen Echolotung (1 bis 20 MHz) gut nach oben bündelnde Richtantennen sehr teuer. Deshalb besteht Interesse an Methoden mit geringem Aufwand.

Eine solche Methode wurde vor einigen Jahren vom Verfasser angegeben [7] und seitdem mit Erfolg benutzt. Ähnlich, wie es in der Funknavigation üblich ist, stellt man durch Umtasten zwischen zwei Diagrammen einen „Leitstrahl“ her. Weil aber zur Aussendung einer festen Richtung im Raum bei Abheben auf Amplitudengleichheit drei Antennen erforderlich wären, wurde ein vereinfachtes Verfahren benutzt, bei dem man auf Verschwinden einer Amplitude abhebt. Auch wegen der komplizierten Fading-Erscheinungen wäre Amplituden-

gleichheit als Kriterium für die Raumwelle möglicherweise unbrauchbar. Man braucht ein Antennendiagramm mit vertikaler Nullstelle. Die Antennen bestehen aus zwei vertikalen Rhombusantennen, die so miteinander verbunden werden, daß die Antennen abwechselnd als vertikaler Rhombus mit Maximum nach oben und als vertikaler Dipol mit Nullstelle nach oben wirksam sind (Bild 5). Bei dieser Anordnung werden die gewünschten vertikalen Echos im Umtastrhythmus unterbrochen, da sie in der zweiten Schaltstellung, der Nullstelle wegen, nicht aufgenommen werden können. Echos aus schräger Richtung zeigen dagegen einen geringeren Umtasteinfluß, bei einem Empfänger mit logarithmischer Empfindlichkeitskurve werden sie praktisch stetig durchgezeichnet. Einige Beispiele aus zahlreichen Registrierungen zeigen die Aufnahmen von Bild 6.

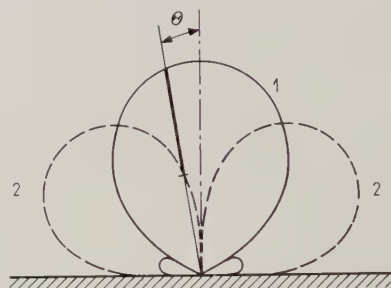


Bild 5. Antennenumtastung zur Unterscheidung von Echos aus vertikaler Richtung.

Wir haben durch diese Registrierungen zwei wichtige Hinweise zur Interpretation der Ionosphärenregistrierungen erhalten: So ergab sich, daß die Echos Spuren von der sogenannten Es-Schicht immer unterbrochen waren, also vertikal einfallen (Bild 6a). Andererseits konnten bei den gestreuten F-Echos zwei Arten unterschieden werden; Fälle, bei denen der diffuse Anteil schräg einfällt (Bild 6b), aber auch andere, bei denen auch der diffuse Anteil vertikal einfällt (Bild 6c). Schließlich stellte sich heraus, daß bei einer Erscheinung, die nicht ganz selten ist, eine Spur mit Sicherheit schräg einfällt; es ist die obere der verdoppelten Echos Spuren, die in Bild 6d zu sehen sind.

Im Abschnitt 3.4 werden wir diskutieren, inwieweit eine Unterscheidung der Einfallrichtung zur Empfangsverbesserung angewendet werden kann.

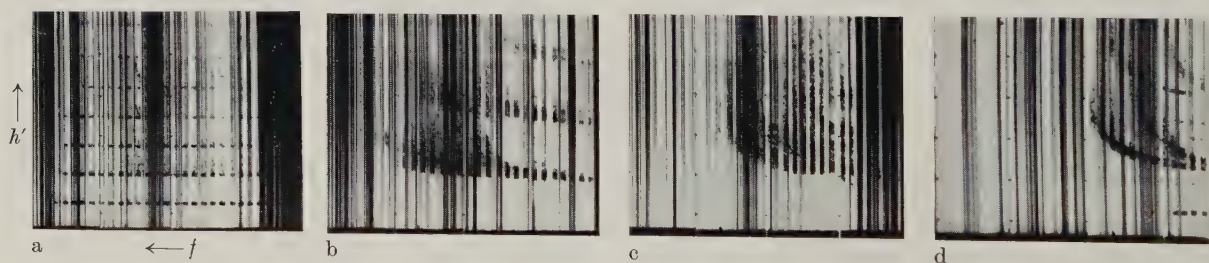


Bild 6. Beispiele von Ionogrammen mit Antennenumtastung.

3. Einige Verfahren zur Empfangverbesserung

Die in den beiden vorausgegangenen Abschnitten diskutierten Kriterien können am einfachsten mit einer auf das Kriterium speziell zugeschnittenen Registrierart herangezogen werden. Während das auf die Periodizität (Abschnitt 1.1) zugeschnittene Verfahren allgemein akzeptiert werden kann, möchte man die speziellen Registrierformen, die wir bei den anderen Registrierverfahren anführten, nicht allgemein benutzen. Die meisten dieser Verfahren bedeuten ja gleichzeitig einen gewissen Informationsverlust, den man ungern in Kauf nehmen wird. Es stellt sich deshalb die Frage, ob nicht eine Methode gefunden werden kann, bei der diese Kriterien lediglich zur Kontrolle der Wiedergabe und gegebenenfalls zur Auslese von Zeichen benutzt werden, ohne daß ein spezielles Registrierverfahren eingeführt würde. Grundsätzlich kann man das mit Hilfe des Mehrkanalempfangs erreichen. Wir besprechen in den folgenden Abschnitten 3.2 bis 3.4 einige dieser Möglichkeiten.

3.1. Impulswiederholung

Durch Impulswiederholung wächst die Empfindlichkeit für die gewünschten Signale, weil dabei effektiv ein Mittelwert mehrerer Kippungen (Anzahl T/τ) registriert wird. Sind U_S und U_N momentane Effektivwerte der Signal- und Störspannung, so addieren sich beide bei zufälliger Phasenlage im Mittel zur resultierenden Spannung

$$\overline{U_{SN}} = \sqrt{\overline{U_S^2} + \overline{U_N^2}}.$$

Das Mittel über die Zeit t (durch Überstreichen angedeutet) ist dann an einer Stelle mit Echo $\overline{U_{SN}}$, dagegen an einer Stelle ohne Echo lediglich $\overline{U_N}$. Früher nahm man häufig vereinfachend an, die günstigsten Empfangsbedingungen seien dann erreicht, wenn die Nutzspannung möglichst hoch über der Störspannung liege. Nach der modernen Fassung der Informationstheorie [8] ist diese Vorstellung jedoch irrig. Es geht um die Möglichkeit, ein Signal trotz des unkontrollierbaren Störpegels als solches zu erkennen. Das ist ein Entscheidungsakt, der von einem menschlichen Beobachter oder auch von einem Automaten ausgeführt werden kann. In jedem Fall wird die Entscheidung nach bestimmten, beim Menschen häufig recht komplizierten Regeln ausgeführt. Bei der Entscheidung wird die a-priori-Kenntnis des unverfälschten Signals mitbenutzt. Man kann den Entscheidungsakt beschreiben als Einteilung des „Beobachtungsraumes“ der möglichen Ausgangswerte in einen „Ja“- und in einen „Nein“-Bereich. Weil der Störpegel unkontrollierbar ist, sind dabei Fehlentscheidungen grundsätzlich unvermeidlich. Es kann passieren, daß ein Störzeichen als Echo identifiziert wird oder auch umgekehrt ein Echozeichen übersehen und dem Störpegel zugeordnet wird. Es zeigt sich nun, daß es für unser Problem keine absolut beste Lösung gibt, höchstens eine „relativ beste“ Lösung, die aber von den willkürlichen Forderungen

des „Kunden“ abhängt. Der „Kunde“, nämlich der Benutzer der Ergebnisse, kann zwar nicht verlangen, daß Fehlentscheidungen ganz ausgeschlossen sind, aber er wird sie nur in beschränktem Maße zulassen. Im folgenden werden wir den Verlust eines Echozeichens im Störpegel als weniger schlimm ansehen, als „falschen Alarm“, nämlich die Verwechslung eines Stör- mit einem Echozeichen.

Wir erkennen ein Echo dann als solches an, wenn das Verhältnis $\overline{U_{SN}}/\overline{U_N}$ sich signifikativ von eins unterscheidet. Setzt man durch Konvention ein Maß für die statistische Sicherheit der Entscheidung fest, so kann nach den Regeln der Statistik die relative Signalamplitude angegeben werden, bei der gerade die Sicherheitsgrenze erreicht wird. Dieser Wert hängt vom Verhältnis T/τ ab, das zugleich die Wiederholungszahl n ist. Je größer die Wiederholungszahl, desto geringere Zeichenamplituden lassen sich noch vom Störpegel unterscheiden. Mit häufiger Wiederholung können noch Echos festgestellt werden, die erheblich unter dem Störpegel liegen. Jedoch ist praktisch die Möglichkeit, durch Wiederholung Information zu gewinnen, dadurch begrenzt, daß der Zustand der Ionosphäre sich laufend ändert und so die mögliche Beobachtungsdauer beschränkt ist.

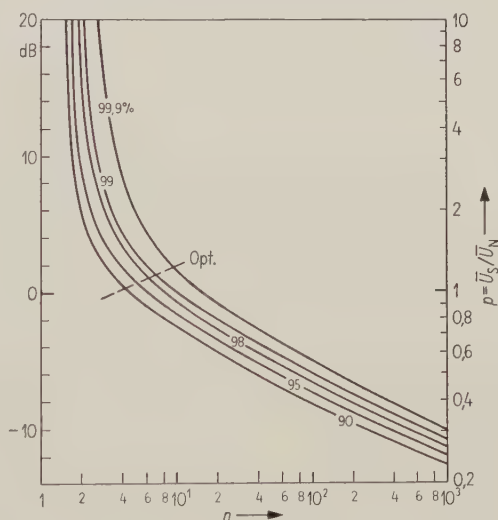


Bild 7. Zusammenhang zwischen Mindestamplitude der Echos über dem Störpegel, p , und Anzahl der Wiederholungen, n , mit verschiedenen, konventionellen Sicherheitswerten als Parameter.

Um sinnvolle Angaben über den Zusammenhang von erforderlicher Amplitude $p = \overline{U_S}/\overline{U_N}$ und Wiederholungszahl n zu machen, muß die Statistik des Störpegels bekannt sein. Sie entspricht bei Ionosphärenlotungen nie der sogenannten „Normalverteilung“ der Energie, wie man sie beim thermischen Rauschen hat. Vielmehr begrenzen praktisch fremde Sendungen oder Atmospherics die Empfangsmöglichkeiten. In beiden Fällen ist die Häufigkeit stärkerer Störsignale größer als bei der „Normalverteilung“. Eine annehmbare Näherung ist eine Log-Normalverteilung der Störspannung, wobei die

mittlere Streubreite mindestens ± 3 dB beträgt (d. h. die Störspannung schwankt im Mittel um $\pm 40\%$).

Unter dieser Annahme ist in Bild 7 der Zusammenhang zwischen p und n wiedergegeben, wobei verschiedene konventionelle Werte der Sicherheit angenommen wurden. Die erforderliche Zeichenamplitude ist für $n = 2$ noch groß, sinkt dann rasch mit wachsendem n , merklich bis $n = 10$.

Für die technische Anwendung wird die Anlage durch eine Nebenbedingung etwas kompliziert. Bei großem Aufwand ist nämlich durch die Röhrenbestückung, insbesondere das Impulsthyratron, die mittlere Leistung des Impulssenders festgelegt; das bedeutet, daß $n \cdot U_s^2$ konstant zu halten ist. Mit dieser Nebenbedingung erhalten wir einen anderen Zusammenhang, der in Bild 7 gestrichelt eingetragen ist. So ergibt sich ein Optimum bei geringer Wiederholungszahl und entsprechend hoher Leistung des Einzelimpulses.

3.2. Differenzieren

„Differenzieren“ bedeutet Beseitigung der niederen Tonfrequenzen nach der Gleichrichtung. Das kann durch irgendeinen Hochpaß geschehen, am einfachsten durch Kopplung mit einem genügend kleinen Kondensator und entsprechendem Ableitwiderstand. Wie im Abschnitt 2.2.2 erläutert, geht beim Differenzieren ein großer Teil der Information verloren. Andererseits steigt dadurch die Empfindlichkeit für Impulse im Vergleich zu Dauerstrichstörungen. Man kann jedoch, wenn man den Empfangskanal gabelt, erreichen, daß lediglich in einem Kontrollkanal der Trick des Differenzierens angewandt wird, nicht aber im eigentlichen Registrierkanal. Die Ausgangsspannung des Kontrollkanals kann nämlich dazu benutzt werden, den Registrierkanal nur dann für eine Zeit lang einzuschalten, wenn im Kontrollkanal ein genügend starkes Zeichen erhalten wurde. Solange der Registrierkanal überhaupt eingeschaltet ist, wird seine Wiedergabe tadellos sein.

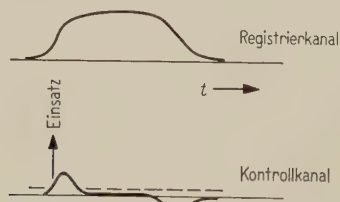


Bild 8. Zeichenform im Kontroll- und im Registrierkanal.

Die praktische Durchführung ist etwas kompliziert. Die beiden Kanäle gabeln sich erst nach der Gleichrichtung der Zwischenfrequenz. Im Kontrollkanal wird das gleichgerichtete Empfangsgut über einen Hochpaß gegeben und nochmals verstärkt, man erhält dadurch zu Beginn jedes Echoimpulses eine positive, zu Ende eine negative Spitze (Bild 8). Mit den positiven Spitzen, die einen einstellbaren Schwellenwert überschreiten, wird dann ein mono-

stabiler Multivibrator angestoßen, der für ein bis zwei Impulsdauern eine positive Rechteckspannung liefert. Im anderen, dem Registrierkanal liegt eine Verstärkerstufe, die von der Spannung des Multivibrators kontrolliert wird und somit jeweils nur für eine Dauer von etwa zwei Impulslängen geöffnet wird.

Die angegebene Schaltung hat noch den Nachteil einer gewissen Verzögerung, weil ja die Front des Impulses schon eine gewisse Steilheit erreicht haben muß, bis der Registrierkanal geöffnet wird. Der Mißstand kann mit dem seinerzeit von HABANN [9] zur Unterdrückung atmosphärischer Störimpulse eingeführten Trick abgestellt werden. Man gabelt schon im Zwischenfrequenzteil und gibt dem Registrierkanal eine Zwischenfrequenzstufe mehr. Die Verzögerung in dieser Stufe reicht aus, um den Mißstand zu beseitigen; die Öffnung des Registrierkanals erfolgt dann sogar kurz vor Beginn des Zeichens in diesem Kanal.

3.3. Hochfrequente Phase

Die hochfrequente Phase kann in ähnlicher Weise zur Kontrolle der Registrierung benutzt werden. Allerdings muß in diesem Fall etwas größerer Aufwand getrieben werden. Erstens ist die Gabelung früher durchzuführen, jedenfalls von der Stufe an, in der die Spannung des Hilfsoszillators dem Kontrollkanal zugemischt wird. Zweitens muß zwischen Kontrolle und Registrierung jetzt eine Zeitdifferenz von etwa einer Impulsdauer eingefügt werden; es kann ja nicht mehr am Anfang des modulierten Zeichens schon sicher entschieden werden, daß diese Modulation kohärent bleibt. Für Ionosphärenregistrierungen beträgt diese Zeitdauer etwa 0,1 ms; eine so lange Zeit kann mit direkten Verzögerungsanordnungen nur schwer ohne wesentliche Verzerrung überbrückt werden. Bei einem einzelnen Impuls müßte man ein Kurzzeit-Gedächtnis anwenden. Verschiedenste Ausführungsformen sind denkbar: Magnettonband, Magnettrommel, elektrostatisches Gedächtnis, akustische Verzögerungsleitung; alle sind ziemlich aufwendig. Bei Impulswiederholung wird man aber an Stelle der Registrierspannung einfacher die Kontrollspannung verzögern. Mit einer Multivibratoranordnung ist es leicht möglich, um fast eine volle Wiederholungsperiode τ zu verzögern.

Nachdem wir in Bild 4 sahen, daß die hochfrequente Phase a priori unbekannt ist und außerdem sich ziemlich rasch verändert, kann in der Kontrollanordnung nur auf die Modulationsfrequenz abgehoben werden. Man wird deshalb in den Kontrollzweig ein sehr schmales, auf die Modulationsfrequenz abgestimmtes Filter einsetzen und damit nur die Zeiten auslesen, zu denen die gewünschte Modulationsfrequenz kräftig auftrat. Nach Gleichrichtung können diese Impulse benutzt werden, um je einen monostabilen Multivibrator anzustoßen, der zuerst eine lange negative Phase und dann eine kurze positive haben muß. Damit ist eine Kontrollspannung für die folgende Impulswieder-

holung erreicht. Mit einem einzigen Multivibrator kann die Schaltung nur ein Echo wiedergeben (das direkte Zeichen wird man ausblenden). Sollte mehr erwünscht sein, so können entsprechende Auswahlbereiche durch Suchimpulse festgelegt werden. In jedem von ihnen kann dann ein Echo gesucht und in die folgenden Perioden übertragen werden. Auf diese Weise sollte es möglich sein, die Auslese der wiedergegebenen Empfangssignale mit Hilfe der Phasenkohärenz zu steuern, ohne die Wiedergabe der Zeichen selbst irgendwie zu beeinflussen. Damit würde der Informationsverlust vermieden werden, den wir im Abschnitt 2.3 angegeben hatten.

3.4. Einfallsrichtung der Echos

Die vertikale Einfallsrichtung kann in verschiedener Weise zur Empfindlichkeitsverbesserung benutzt werden. Man benötigt mindestens zwei Antennen und entweder zwei Empfänger oder einen umschaltbaren Empfänger. Bei elektronischer Umschaltung wird man zweckmäßig das Doppel der Wiederholungsperiode τ als Umschaltperiode wählen, und für jede zweite Periode eine andere Antenne benutzen. Man kann dann in ähnlicher Weise, wie im Abschnitt 3.3 beschrieben, eine Periode zur Kontrolle der darauffolgenden benutzen.

Die grundsätzlichen Überlegungen wollen wir für den Fall zweier getrennter Empfänger durchführen. Man könnte daran denken, einfach die Differenz der gleichgerichteten, momentanen Ausgangsspannung zu bilden. Bei logarithmischer Empfindlichkeitskurve würde dadurch ein Amplitudenwert erreicht, der dem dB-Abstand des Zeichens vom Störpegel entspräche. Dadurch ginge aber die Eindeutigkeit der Ausgangsspannung verloren, die Anzeige würde bei hohem Störpegel kleiner als bei geringem erscheinen. Eine Kontrollmethode könnte

auch hier diese Nachteile vermeiden. Man kann etwa die Differenz der Ausgangsspannungen ($U_a - U_b$) als Regelspannung für eine positive Empfindlichkeitsregelung benutzen, oder auch sie bei einer bestimmten Amplitude abschneiden und als Schaltungsspannung verwenden. Bei Amplitudenregistrierung könnte damit etwa die Helligkeit frei gegeben werden, bei Helligkeitsregistrierungen würde man den Empfang nur während der Zeiten freigeben, bei denen die Differenz ausreichend groß ist. Damit würde wiederum die Eindeutigkeit des Ausgangs erhalten bleiben, wenigstens soweit überhaupt Ausgangssignale wiedergegeben werden. Je nach der Anordnung können dabei allerdings sehr kleine Echosignale verloren gehen.

Die angegebenen Methoden sind bisher nur zum kleinen Teil beim Echo-Empfang benutzt worden. Es scheint, daß mit ihrer Anwendung manche Verbesserungen auch der Routine-Beobachtungen möglich wären; jedenfalls geben sie für spezielle Beobachtungen einige interessante Möglichkeiten.

Schrifttum

- [1] BIBL, K., Aktive Hochfrequenzspektrometer für ionosphärische Echolotung; I. Direktregistrierung ionosphärischer Charakteristiken. A. E. Ü. **14** [1960], 341–347.
- [2] HUBER, L. und RAWER, K., Zur Frage des „besten“ Impulsempfängers, Vergleich verschiedener Impulsverstärker. A. E. Ü. **4** [1950], 475–484, 523–526.
- [3] First Report of the Special Committee on World Wide Ionospheric Soundings. URSI Inform. Bull. Nr. 99 [1950], 48–90.
- [4] RANZI, I., Recording wireless echoes at the transmitting station. Nature **132** [1933], 174.
- [5] FINDLAY, J. W., The phase and group paths of radio waves returned from region E of the ionosphere. J. atmos. terrest. Phys. **1** [1951], 353–366.
- [6] MC NICOL, R. W. E., WEBSTER, H. C. und BOWMAN, G. G., A study of „spread-F“ ionospheric echoes at night at Brisbane. Austral. J. Phys. **9** [1956], 247–271.
- [7] BIBL, K., HARNISCHMACHER, E. und RAWER, K., Some observations of ionospheric movements. „The physics of the ionosphere“, London 1955, 114–118.
- [8] VAN METER, D. und MIDDLETON, D., Modern statistical approach to reception in communication theory. Transact. Inst. Radio Engrs. PPT [1954], 119–145.
- [9] HABANN, Unveröffentlichter Bericht, 1943.

BUCHBESPRECHUNG

W. C. Hoffman, *Statistical methods in radio wave propagation* (Proceedings of a symposium held at the University of California, Los Angeles, 18. bis 20. Juni 1958). Pergamon Press Ltd., London 1960, XIV, 334 Seiten, zahlr. Bilder, 19 cm × 25 cm, Ganzleinen sh 90/—.

Jeder Wissenschaftler der Sektion Wellenausbreitung in Troposphäre und Ionosphäre wird das vorliegende Buch dankbar begrüßen; füllt es doch eine wesentliche Lücke in der Fachliteratur. Niemand wird es wohl bedauern, daß sich hinter dem Titel kein Lehrbuch verbirgt, sondern eine Sammlung von 23 Vorträgen, die während eines gleichlautenden Symposiums in Los Angeles im Juni 1958 gehalten wurden.

10 Vorträge machen den ersten Abschnitt des Buches aus. Wie dessen Titel „Statistik und Methodik“ schon sagt, kommt darin vorwiegend die abstrakte Theorie zu Wort. Die Vorträge selbst lassen jedoch erkennen, daß auch ihnen immer ein spezielles physikalisches Problem zugrunde liegt. Der

zweite Abschnitt des Buches umfaßt 11 Vorträge. Sie befassen sich ausschließlich mit der statistischen Auswertung troposphärischer und ionosphärischer Daten. Die beiden letzten Vorträge, und damit der 3. Abschnitt des Buches, befassen sich mit der speziellen Instrumentation für statistische Untersuchungen von Meßdaten aus der Wellenausbreitung.

Die enge Verbindung zwischen Theorie und Praxis macht das Buch besonders wertvoll. Gelegentliche Wiederholungen sind bei einem derartigen Werk unvermeidlich; sie schaden aber keineswegs. Im Grunde kann ein derartiges Buch auch kein abgeschlossenes Ganzes bilden. Diesen Nachteil nahm der Herausgeber in Kauf, um dafür möglichst schnell die Lücke in der einschlägigen Fachliteratur zu schließen. Wo Fehlendes nachgelesen werden kann, vermerkt der Herausgeber aber immer durch eingefügte Literaturhinweise.

Auch die äußere Ausführung des Buches entspricht dem hohen Standard des Werkes. Zusammenfassend kann man wohl sagen: Dem Buch wird eine dankbare Aufnahme sicher sein.

W. BECKER

Oberflächenströme in Inversionsschichten an Halbleitern

Von EBERHARD GROSCHWITZ, ERNST HOFMEISTER und RUDOLF EBHARDT

Mitteilung aus dem Wernerwerk für Bauelemente der Siemens & Halske AG, München

(A.E.Ü. 14 [1960], 380—396; eingegangen am 15. Juli 1960)

DK 621.382.2

Die vorliegende Arbeit bringt eine Darstellung einer Reihe von theoretischen und experimentellen Ergebnissen über Oberflächenströme in Inversionsschichten an Halbleiteroberflächen. Es werden die Struktur der Inversionsschichten und die physikalischen Vorgänge, die zu Strukturänderungen führen, beschrieben. Der Strommechanismus in Inversionsschichten bei Anlegen eines äußeren Potentials wird theoretisch am Beispiel von Spitzenkontakt-Gleichrichtern abgeleitet. Experimentelle Ergebnisse über die Kennlinien-Eigenschaften von formierten Spitzenkontakt-Gleichrichtern werden mitgeteilt. Eine Deutung der gemessenen Werte durch die analytisch darstellbaren Oberflächenströme ist in befriedigender Weise möglich. Der Einfluß der Halbleiteroberfläche und der Gestalt des formierten Bereichs auf den Gleichrichtereffekt wird quantitativ untersucht. Es zeigt sich, daß bei Spitzenkontakt-Gleichrichtern der Oberflächenstrom sowohl die Größe des Diodenstromes als auch die Gestalt der Kennlinie entscheidend bestimmt. Es werden zunächst nur die Verhältnisse bei relativ kleinen Gleichspannungen untersucht, damit der Einfluß der Oberflächeneffekte nicht durch Nebenerscheinungen überdeckt wird.

This paper presents a number of theoretical and experimental results concerning surface currents in inversion layers of semiconductors, describing the structure of the inversion layers and the physical phenomena giving rise to structural changes. The current mechanism taking place in inversion layers with application of an external potential is derived theoretically by reference to the example of point contact rectifiers. Experimental results concerning the properties of the characteristic curves of formed point contact rectifiers are communicated. The measured values can be interpreted satisfactorily by the surface currents which can be presented analytically. The influence on the rectification effect of the semiconductor surface and the shape of the formed zone is investigated quantitatively. It turns out that with point contact rectifiers the surface current determines in a decisive manner both the magnitude of the diode current and the shape of the characteristic curve. The investigation relates initially merely to the conditions encountered with relatively low direct voltages to prevent secondary phenomena from masking the influence of the surface effects.

1. Einleitung

Oberflächenströme spielen in der Technik der Halbleiterbauelemente eine wesentliche, meist unerwünschte Rolle. Infolge von hierdurch bedingten Nebenschlüssen treten an stromdurchflossenen p-n-Übergängen Anomalien auf, die sich durch erhöhte Sperr- und Flußströme sowie zeitliche Inkonzanz der Stromspannungscharakteristik und vermehrtes Rauschen bemerkbar machen. Diese Oberflächenströme liefern häufig einen Beitrag zum Gesamtstrom von gleicher Größenordnung wie der Volumenstrom. Hierdurch werden die elektrischen Eigenschaften gleicher Bauelemente bei jedem einzelnen Exemplar mehr oder weniger stark modifiziert. Aus diesen Gründen ist es zweckmäßig, die mit der Existenz von Oberflächenströmen in Zusammenhang stehenden Erscheinungen nicht als unerwünschte Nebeneffekte zu betrachten, sondern den Vorgängen an der Kristalloberfläche und in der Oberflächeninversionsschicht als selbständigem Fragenkomplex ein besonderes Interesse zuzuwenden. Man darf vermuten, daß Oberflächenströme in Halbleitern in Zukunft auch technische Nutzanwendung finden werden.

Man unterscheidet Transportvorgänge von Ladungsträgern innerhalb der adsorbierten Fremdschicht auf der Kristalloberfläche von Strömen in der Oberflächeninversionsschicht. Bei starker Inversion können Strombeiträge in adsorbierten

Fremdschichten von monomolekularer Schichtdicke im allgemeinen gegen Oberflächenströme in Inversionsschichten vernachlässigt werden. Bei Halbleiteroberflächen ohne stark invertierte Anreicherungsschichten können jedoch auch Ströme in adsorbierten Fremdschichten eine Rolle spielen [1], [2].

Die vorliegende Arbeit handelt von Oberflächenströmen in Inversionsschichten bei starker Inversion. Diese Verhältnisse sind an Germanium und Silizium realisierbar und in den meisten Fällen von selbst gegeben. Untersuchungen der physikalischen Verhältnisse in solchen Kanälen sowie Messungen der Leitfähigkeit wurden insbesondere von J. BARDEEN und W. H. BRATTAIN [3] und anderen Autoren durchgeführt [4]—[10]. Untersuchungen von Oberflächeninversionsschichten wurden zuerst von A. I. GUBANOW [11] und später von anderen unternommen [12], [13]. Eine ausführliche Darstellung des Problemkreises stammt von H. U. HARTEN und W. SCHULTZ [14].

Die Diskussion der im folgenden mitgeteilten experimentellen Ergebnisse basiert auf einer analytischen Beschreibung der Oberflächenströme, die einer der Verfasser gegeben hat. Wesentlich hierbei ist, in welcher Weise die Beschaffenheit der Kristalloberfläche, die Parameter des Halbleiters sowie Rekombinations- und Paarbildungsprozesse und Diffusions-effekte die Strom-Spannungscharakteristik beeinflussen. Das Wechselstromproblem wird in einer späteren Arbeit behandelt.

Die theoretischen Ergebnisse gestatten eine Analyse neuer Messungen über die Strom-Spannungsabhängigkeit von formierten Spitzkristallgleichrichtern. Die Konzeption zur Beschreibung der Oberflächenströme ist so allgemein gehalten, daß sie grundsätzlich auch auf andere Halbleiterbauelemente angewendet werden kann. Bei Spitzendioden ist jedoch das Problem der Oberflächeneinflüsse viel gravierender als bei Flächendioden oder -transistoren, weil bei Spitzendioden die Ausdehnung der durch die Feder gebildeten Kontaktfläche außerordentlich viel kleiner ist als die geometrische Größe der physikalisch wirksamen Halbleiteroberfläche [15], [16]. Aus diesem Grunde stellen Spitzkontakt-Gleichrichter eine physikalische Anordnung dar, an der sich der Einfluß und die grundsätzlichen Wesenszüge von Oberflächenströmen besonders gut studieren lassen.

Die analytischen Ergebnisse, mit denen die experimentellen Kurven verglichen werden, sind im Anhang zusammengestellt.

2. Aufbau der Oberflächeninversionsschicht

2.1. Inversionsschicht ohne äußeres Potential

Struktur und Aufbau einer Oberflächeninversionsschicht eines beispielsweise n-leitenden Halbleiterkristalls sind schematisch in Bild 1 dargestellt. Das Diagramm zeigt den Verlauf der Konzentration der beweglichen Ladungsträger und des elektrostatischen Potentials in Abhängigkeit von der Kristalltiefe, gemessen vom Kristallrand aus. Die Länge l bezeichnet die Tiefe der gesamten Raumladungszone bis zur Grenze des neutralen Gebietes im Halbleiterinneren. Die Inversionsebene befindet sich an der Stelle x_1 , wo ersichtlich die Konzentrationen der Elektronen n und der Löcher p gleich der Eigenleitungsdichte n_i sind. Die Dicke x' der randnahen Anreicherungsschicht begrenzt eine Zone, in der bei starker Inversion die im Kristallinneren als Minoritätsträger vorhandenen Defektelektronen eine we-

sentlich höhere Konzentration annehmen als die Donatorenkonzentration n_D . Für die Existenz von Oberflächenströmen spielt deshalb diese Anreicherungsschicht eine entscheidend wichtige Rolle.

Eine Oberflächeninversionsschicht unterscheidet sich in wesentlichen Eigenschaften von einem konventionellen p-n-Übergang im Kristallinneren. Die Umkehr des Leitungstypus erfolgt bei einer Oberflächeninversionsschicht nicht durch einen Wechsel der Dotierung, sondern wird durch eine infolge von Adsorption und Oberflächenzuständen an der Kristalloberfläche vorhandene elektrische Flächenladungsdichte erzeugt. Der Potentialverlauf hat bei einer Oberflächeninversionsschicht infolge der sich bis zur Kristalloberfläche erstreckenden gleichartigen Dotierung keinen Wendepunkt, wie dies bei einem p-n-Übergang im Inneren des Kristalls wegen des Wechsels der Dotierung der Fall ist. Infolgedessen findet in einer Inversionsschicht an der Halbleiteroberfläche eine ständig zunehmende Abkrümmung des Potentials bis zum Kristallrand statt.

Die elektrische Flächenladungsdichte an der Kristalloberfläche ist nach einem allgemeinen Satz der Elektrostatik proportional zur Randfeldstärke. Insofern wird durch die elektrische Flächenladungsdichte als Randwert für den Potentialverlauf der Oberflächeninversionsschicht die elektrische Randfeldstärke festgelegt. Die elektrische Flächenladungsdichte hängt von der Vorgeschichte und der hierbei erzeugten molekularen Struktur der Kristalloberfläche ab und kann durch entsprechende Behandlung variiert werden. Die Elementarladungen an der Oberfläche verteilen sich auf ein System von Energiezuständen. Man konnte experimentell feststellen, daß offenbar zwei verschiedene Arten von Zentren existieren, die sich durch ihre Einstellzeiten unterscheiden [7]—[9]. Diejenigen mit einer Relaxationszeit von 10^{-7} bis 10^{-8} s werden als „schnelle“ und die anderen mit Relaxationszeiten in der Größenordnung von Sekunden bis Stunden als „langsame“ Oberflächenzustände bezeichnet. Die effektive Zustandsdichte der „schnellen“ Zustände beträgt ungefähr 10^{11} cm $^{-2}$; hierbei handelt es sich um vorwiegend akzeptorielle Terme in der Nähe des Valenzbandes. „Schnelle“ Zustände mit Donatorencharakter kommen auch vor. Deren effektive Zustandsdichte ist aber häufig geringer und liegt bei etwa 10^{10} cm $^{-2}$. Die „schnellen“ Zentren spielen insbesondere bei elektrischen Wechselvorgängen eine Rolle, deren Periode mit den zugehörigen Relaxationszeiten vergleichbar ist. Für den Aufbau der Oberflächeninversionsschicht sind jedoch die „langsamen“ Zentren entscheidend. Ihre effektive Zustandsdichte kann Werte zwischen 10^{11} cm $^{-2}$ und 10^{15} cm $^{-2}$ annehmen. Bei der Verteilung der Elementarladungen auf diese Zustände liegen die Werte der elektrischen Randfeldstärke im allgemeinen in einem Bereich zwischen 10^3 V/cm und 10^6 V/cm. Die Randfeldstärke bzw. ihre entsprechende effektive, elektrische Flächenladungsdichte stellt eine charakteristische, phänomenologische Größe dar, durch die summarisch jene

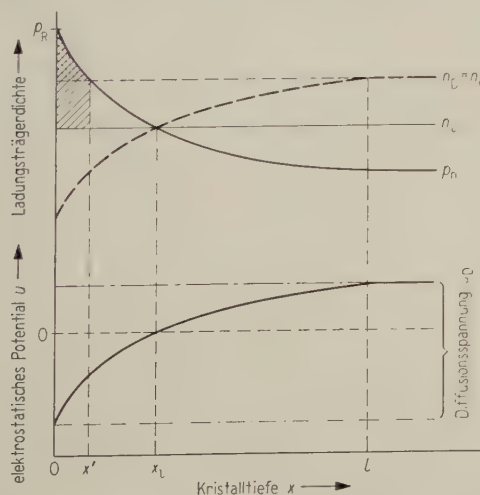


Bild 1. Schema der Konzentrationsverläufe und des dimensionslosen elektrostatischen Potentials $u = q\psi/kT$ (Voltäquivalent) in Oberflächen-Inversionsschichten (Ordinate der Konzentrationsverläufe im logarithmischen Maßstab).

Eigenschaften der Kristalloberfläche gekennzeichnet werden, die den Aufbau der Oberflächeninversionsschicht vom Rand her bestimmen. Als weitere, die Struktur der Oberflächeninversionsschicht beeinflussende Parameter sind die Dotierung bzw. der spezifische Widerstand, die Eigenleitungsdichte, die Dielektrizitätskonstante und die Temperatur zu nennen.

Der Zusammenhang zwischen Randfeldstärke und der effektiven, elektrischen Flächenladungsdichte an der Halbleiteroberfläche ist in Bild 2 dargestellt. Den Potentialverlauf innerhalb der Oberflächeninversionsschicht bei starker Inversion zeigt Bild 3. Im Gegensatz zu Bild 1 handelt es sich in diesem Diagramm um berechnete Potentialverläufe bei verschiedenen Randfeldstärken und Dotierungen [17]. Das elektrostatische Potential ψ ist so gewählt, daß bei $x = x_1$ (Inversionsebene) das aufgetragene dimensionslose Potential (Voltäquivalent) $u = q\psi/kT = 0$ ist. Der sich hierbei er-

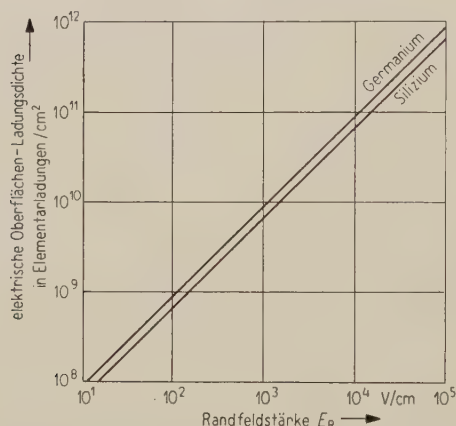


Bild 2. Zusammenhang zwischen Randfeldstärke und Oberflächenladung.

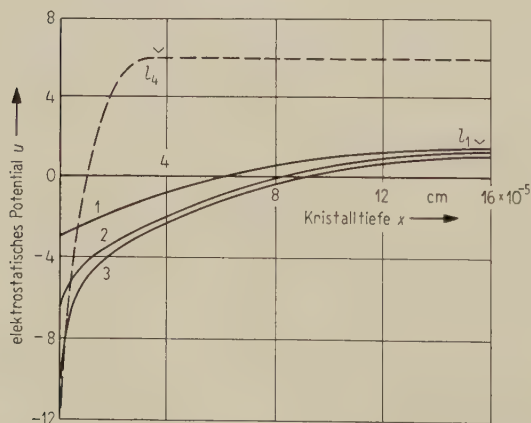


Bild 3. Berechneter Verlauf des dimensionslosen elektrostatischen Potentials $u = q\psi/kT$ in einer Inversionsschicht bei starker Inversion. l : Tiefe der Raumladungszone, n-Germanium, $T = 300^\circ\text{K}$.

Kurve 1: $E_R = 2 \cdot 10^3 \text{ V/cm}$, $n_D = 10^{14} \text{ cm}^{-3}$
bzw. $\rho = 15 \Omega \text{ cm}$,

Kurve 2: $E_R = 1 \cdot 10^4 \text{ V/cm}$, $n_D = 10^{14} \text{ cm}^{-3}$,

Kurve 3: $E_R = 1 \cdot 10^5 \text{ V/cm}$, $n_D = 10^{14} \text{ cm}^{-3}$,

Kurve 4: $E_R = 1 \cdot 10^5 \text{ V/cm}$, $n_D = 10^{16} \text{ cm}^{-3}$
bzw. $\rho = 0,16 \Omega \text{ cm}$.

gebende gesamte Potentialabfall (Diffusionspotential u_D) innerhalb der Raumladungszone sowie deren räumliche Erstreckung, gemessen vom Kristallrand aus, sind in Einklang mit experimentellen Ergebnissen [3]–[9]. Die Raumladungstiefen der Kurven 2 und 3 sind nicht markiert, weil sie außerhalb des Diagrammes liegen würden. Mit wachsender Donatorkonzentration und somit bei kleinerem spezifischen Widerstand des Halbleiters wird die Tiefe der Raumladungszone immer kleiner. Der Grund hierfür ist, daß die elektrische Ladung der Oberflächenbelegung mit ihren Feldlinien in den Halbleiter tiefer hineinwirkt, wenn seine Leitfähigkeit klein ist. Diese Gesetzmäßigkeiten sind in Bild 4 anschaulich gemacht. Es ist das Verhältnis der Defektelektronenkonzentration zur Donatorkonzentration in Abhängigkeit von der Kristalltiefe für verschiedene Werte des spezifischen Widerstandes eines n-leitenden Germaniumkristalls bei einem konventionellen Wert der Randfeldstärke 10^5 V/cm dargestellt. Die Kurvenverläufe sind theoretisch berechnet [17]. Die randnahe Anreicherungszone ist jeweils durch den Teil des Konzentrationsverlaufes gekennzeichnet, bei dem das Verhältnis p/n_D größer als eins ist.

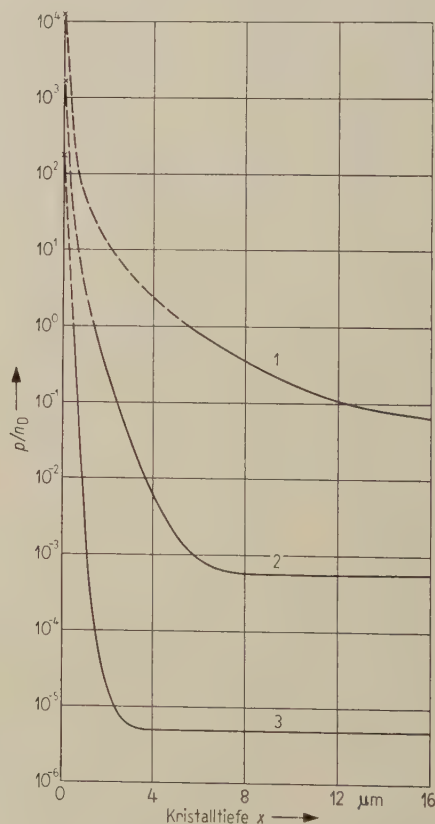


Bild 4. Relative Dichteverteilung der Defektelektronen in Abhängigkeit von der Kristalltiefe bei verschiedenem spezifischem Widerstand eines n-leitenden Halbleiter-Kristalls (berechnete Kurven). Der gestrichelte Kurventeil kennzeichnet das Anreicherungsgebiet starker Inversion; $E_R = 10^5 \text{ V/cm}$ bzw. $Q = 8,8 \cdot 10^{11} \text{ Elementarladungen/cm}^2$.

Kurve 1: $n_D = 10^{14} \text{ cm}^{-3}$ bzw. $\rho = 15 \Omega \text{ cm}$,

Kurve 2: $n_D = 10^{15} \text{ cm}^{-3}$ bzw. $\rho = 1,5 \Omega \text{ cm}$,

Kurve 3: $n_D = 10^{16} \text{ cm}^{-3}$ bzw. $\rho = 0,16 \Omega \text{ cm}$.

2.2. Beeinflussung der Inversionsschicht durch ein äußeres Potential

Die Konzentrationsverteilung der beweglichen Ladungsträger und der zugehörige Verlauf des elektrostatischen Potentials in der Raumladungszone können auf verschiedene Weise geändert werden.

Einmal können die atomaren Stoffparameter des Halbleiters und der Oberfläche modifiziert werden. Dieser Einfluß auf den Verlauf des elektrostatischen Potentials durch Änderung der Dotierung oder der Randfeldstärke bzw. der elektrischen Oberflächenladungsdichte, z. B. hervorgerufen durch Wechsel von adsorbierten Oberflächenschichten, ist in Bild 3 an Hand von berechneten Kurven dargestellt.

Zum anderen kann der Konzentrationsverlauf der beweglichen Ladungsträger im Inneren der Raumladungsschicht durch Einwirkung eines äußeren elektrischen Feldes, das senkrecht zur Kristalloberfläche gerichtet ist, verändert werden. Experimentell wird eine solche Belastung durch ein äußeres Potential im Prinzip entweder durch Annäherung einer geladenen Elektrode an die Halbleiteroberfläche (Bild 7a) oder durch Kontaktierung des Kristalls mit einer äußeren Spannung realisiert. (Bild 7b zeigt ein idealisiertes Schema einer solchen Belastung, wobei der Kontakt homogen über die ganze Oberfläche gedacht ist.) Die zwei Fälle zeichnen sich durch einen grundsätzlichen Unterschied aus.

Bei einer elektrostatischen Belastung durch ein äußeres Potential ist die Inversionsschicht nicht in einen äußeren Stromkreis eingeschlossen. Es findet nur so lange ein Ladungstransport in der Raumladungszone statt, bis sich ihr innerer Aufbau entsprechend den äußeren Bedingungen neu eingestellt hat. Bei Gleichbelastung tritt somit nach sehr kurzer Zeit ein elektrostatischer Gleichgewichtszustand ein, bei dem der durch die Inversionsschicht fließende Gesamtstrom infolge vollständiger Kompensation des Feld- und des Diffusionsstromanteiles verschwindet.

Der andere Fall der Belastung durch ein äußeres Potential ist im Gegensatz hierzu dadurch gekennzeichnet, daß ein Strom senkrecht durch die Oberflächeninversionsschicht fließt, die durch entsprechende Kontaktierung des Kristalls nunmehr in einen Stromkreis geschaltet ist (Bild 7b). Hierbei wird der innere Aufbau der Raumladungszone ebenfalls modifiziert. Bei Gleichbelastung entsteht jedoch jetzt ein stationärer Nichtgleichgewichtszustand, der formal auch als dynamischer Gleichgewichtszustand betrachtet werden kann. Der Feld- und der Diffusionsstromanteil kompensieren sich hierbei nicht vollständig, was einen resultierenden senkrecht zur Inversionsschicht fließenden Gesamtstrom zur Folge hat.

Sowohl bei rein elektrostatischer Belastung der Oberflächeninversionsschicht als auch im Fall einer strombehafteten Belastung wird das elektrostatische Potential im Inneren der Raumladungszone und des Halbleiters gegen den ursprünglichen Verlauf ohne äußeres Potential angehoben oder abgesenkt (Bild 5), wodurch sich auch der hierdurch kausal bedingte

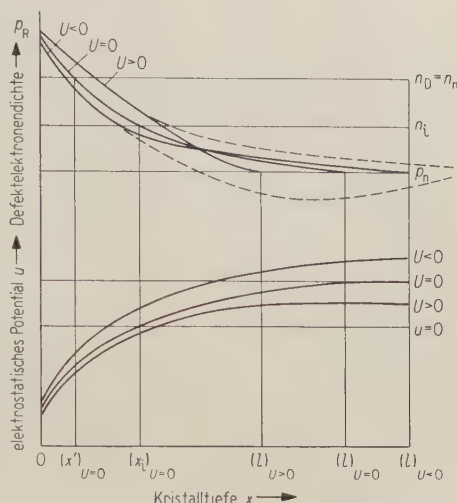


Bild 5. Schematische Darstellung des Konzentrations- und Potentialverlaufes in der Randschicht beim Anlegen einer äußeren Spannung senkrecht zur Kristalloberfläche;

ausgezogene Kurven oben: elektrostatische Belastung,

gestrichelte Kurven oben: Strombelastung,

Kurven unten: Potentialverläufe (elektrostatischer Fall, der strombelastete Fall weicht unwesentlich davon ab).

Konzentrationsverlauf der beweglichen Ladungsträger jeweils entsprechend den vorhandenen Bedingungen neu einstellt. Diese Neueinstellung hat andererseits wieder eine Rückwirkung auf die Randfeldstärke und somit auf die elektrische Oberflächenladung zur Folge. Es werden hierbei an der Kristalloberfläche Flächenladungen influenziert, und zwar von innen her, wodurch der Aufbau der Ladungsverteilung in der Raumladungszone mitbestimmt wird. Infolgedessen ändert sich bei Anwesenheit eines äußeren Potentials in der Oberflächeninversionsschicht auch über die Änderung der Randfeldstärke der sich wechselseitig bedingende Potential- und Konzentrationsverlauf. Der Ausgangspunkt für eine quantitative Formulierung dieser Verhältnisse ist die Gleichgewichtsbedingung der Neutralität für die gesamte elektrische Ladung des Systems.

Wir definieren — wie bei gewöhnlichen p-n-Übergängen üblich — eine senkrecht zur Oberfläche angelegte Spannung als Sperrbelastung, wenn der Minuspol auf der Seite der invertierten p-Schicht liegt und die umgekehrte Polung als Flußbelastung.

Die Berechnung zeigt [17] im Einklang mit der Erfahrung [20], daß bei Potentialbelastungen bis etwa ± 50 mV die Randfeldstärke praktisch spannungsunabhängig ist. Bei stärkerer Sperrbelastung nimmt die Randfeldstärke zu, hingegen wird bei Verkleinerung der Sperrbelastung bzw. bei wachsender Flußbelastung die Randfeldstärke verkleinert.

In Bild 5 sind die Änderungen des Konzentrations- und Potentialverlaufes in der Raumladungszone beim Anlegen eines äußeren Potentials schematisch dargestellt. Diese qualitative, schematische

Darstellung wurde gewählt, um die grundsätzlichen Wesenszüge des physikalischen Sachverhaltes stärker hervorzuheben, als diese bei den angelegten kleinen Spannungen in Wirklichkeit quantitativ in Erscheinung treten. Bild 5 zeigt den schematischen Konzentrationsverlauf der Defektelektronen in der Oberflächenraumladungszone eines n-leitenden Kristalls bei relativ kleinen Belastungen im rein elektrostatischen Gleichgewichtsfall, also ohne resultierenden Gesamtstrom. Bemerkenswert hierbei ist, daß die Defektelektronen am Ende der Raumladungszone und damit am Rande $x = l$ der beginnenden quasineutralen Zone auch bei äußerer Belastung ihre thermische Gleichgewichtskonzentration wie im unbelasteten Fall beibehalten. Im Gegensatz hierzu wird im stationären Nichtgleichgewichtszustand bei Strombelastung (Bild 5) die Defektelektronenkonzentration an der Stelle $x = l$ spannungsabhängig angehoben bzw. abgesenkt. Hierdurch entsteht anschließend in der quasineutralen Zone ein Diffusionsschwanz der Minoritätsträger, welcher den senkrecht durch die Oberflächeninversionsschicht fließenden Strom bestimmt.

Für beide Fälle, mit und ohne Strombelastung, gilt die Gesetzmäßigkeit, daß sich die gesamte Raumladungszone l mit wachsender Sperrbelastung ausdehnt, wobei jedoch die Dicke x' der Anreicherungsschicht bzw. die Tiefe x_1 der Inversionsschicht abnimmt. Bei Flußbelastung hingegen wird die Raumladungsschicht l mit größer werdender Spannung dünner, während die Tiefe der Anreicherungsschicht x' bzw. die Dicke x_1 zunimmt. Diese Gesetzmäßigkeit der geometrischen Veränderungen im Aufbau der Oberflächeninversionsschicht bei Einwirkung eines äußeren Potentials ist in Bild 6 nochmals quantitativ an einem berechneten Beispiel eines Germaniumkristalls mit konventionellen physikalischen Parametern wiedergegeben.

Dieser Effekt hat im wesentlichen folgende Ursache: Die Flußspannung setzt Elektronen aus dem quasineutralen Halbleiterinneren in Richtung auf den Kristallrand in Bewegung. Der Elektronenüberschuß neutralisiert einen gewissen Prozentsatz der Donatorenladungen, wird aber auch andererseits in der Inversionsschicht teilweise durch eine Vermehrung der Defektelektronen ausgeglichen. Dadurch sammeln sich Defektelektronen in der Anreicherungsschicht an und verteilen sich tiefer in die Raumladungszone hinein. Das Massenwirkungsgesetz ist hierbei nicht mehr erfüllt. Umgekehrt verursacht eine in Sperrichtung angelegte Spannung eine Abnahme in der Verteilung der Defektelektronen und somit eine entsprechende Verkleinerung der Tiefe der Anreicherungsschicht. Bemerkenswert ist, daß bei einer hier vorliegenden freien Kristalloberfläche im Gegensatz zu einem Metall-Halbleiterkontakt sowohl die Randfeldstärke E_R als auch das Randpotential u_R mit der angelegten äußeren Spannung variiert. Diese in Bild 5 ebenfalls schematisch angedeutete spannungsabhängige Veränderung der Randwerte der Trägerkonzentration und des Potentials bzw. der elektrischen Feldstärke ist jedoch relativ gering im Vergleich zu der Verschiebung dieser Größen in tiefer gelegenen Gebieten der Raumladungszone.

Die Gesetzmäßigkeiten der in den Bildern 5 und 6 dargestellten geometrischen Veränderung des inneren Aufbaues der Raumladungszone infolge eines senkrecht zur Inversionsebene gerichteten äußeren Potentials sind für den weiter unten zu besprechenden Mechanismus des parallel zur Oberfläche in der Anreicherungsschicht fließenden Oberflächenstromes von grundsätzlicher Bedeutung. Hierbei spielen folgende physikalischen Verhältnisse eine Rolle.

Bei hinreichend starker Inversion wird der quantitativ entscheidende Beitrag des Oberflächenstromes durch die beweglichen Ladungsträger in der randnahen Anreicherungsschicht geliefert. Für die Oberflächenleitfähigkeit in dieser Schicht ist effektiv als experimentell in Erscheinung tretende Größe die an jeder Stelle senkrecht zur Oberfläche gemittelte Konzentration der Defektelektronen (bei n-leitendem Kristall) maßgebend. Die Spannungsabhängigkeit dieser mittleren Trägerkonzentration bestimmt den unipolaren Charakter des parallel zur Oberfläche in der Anreicherungsschicht fließenden Oberflächenstromes. Beim Anlegen eines äußeren Potentials an die Oberflächeninversionsschicht senkrecht zur Inversionsebene im Sinne einer wachsenden Flußbelastung nimmt die mittlere Konzentration der Defektelektronen in der Anreicherungsschicht zu. Die Anreicherungsschicht dehnt sich hierbei aus. Umgekehrt verursacht eine in Sperrichtung angelegte Spannung eine Abnahme der mittleren Defektelektronenkonzentration und eine entsprechende Verkleinerung der Tiefe der Anreicherungsschicht.

Für die weiter unten zu erläuternde Konzeption einer quantitativen Beschreibung des in der Anreicherungsschicht parallel zur Oberfläche fließenden Oberflächenstromes ist außerdem noch der folgende physikalische Umstand zu berücksichtigen. Vergleicht man den Konzentrationsverlauf der invertierten Ladungsträger innerhalb der Anreicherungsschicht bei einer senkrecht zur Inversionsebene angelegten äußeren Belastung, so sind die Konzentrationsverteilungen in der Anreicherungsschicht bei rein elektrostatischer Belastung und bei Strombelastung (Bild 5) bei relativ kleinen äußeren

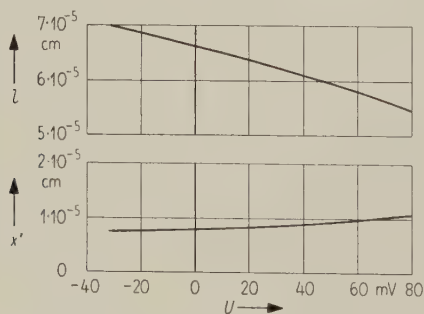


Bild 6. Berechnete Spannungsabhängigkeit der Tiefe x' der Anreicherungsschicht und der Tiefe l der gesamten Raumladungszone; n-Germanium, $T = 300^\circ \text{ K}$, $n_D = 10^{15} \text{ cm}^{-3}$ bzw. $\varrho = 1,5 \Omega \text{ cm}$, $E_R = 10^4 \text{ V/cm}$.

Bei hinreichend starker Inversion wird der quantitativ entscheidende Beitrag des Oberflächenstromes durch die beweglichen Ladungsträger in der randnahen Anreicherungsschicht geliefert. Für die Oberflächenleitfähigkeit in dieser Schicht ist effektiv als experimentell in Erscheinung tretende Größe die an jeder Stelle senkrecht zur Oberfläche gemittelte Konzentration der Defektelektronen (bei n-leitendem Kristall) maßgebend. Die Spannungsabhängigkeit dieser mittleren Trägerkonzentration bestimmt den unipolaren Charakter des parallel zur Oberfläche in der Anreicherungsschicht fließenden Oberflächenstromes. Beim Anlegen eines äußeren Potentials an die Oberflächeninversionsschicht senkrecht zur Inversionsebene im Sinne einer wachsenden Flußbelastung nimmt die mittlere Konzentration der Defektelektronen in der Anreicherungsschicht zu. Die Anreicherungsschicht dehnt sich hierbei aus. Umgekehrt verursacht eine in Sperrichtung angelegte Spannung eine Abnahme der mittleren Defektelektronenkonzentration und eine entsprechende Verkleinerung der Tiefe der Anreicherungsschicht.

Spannungen nur wenig voneinander verschieden. Aus diesem Grunde kann in guter Näherung in beiden Fällen mit der gleichen senkrecht zur Oberfläche gemittelten Defektelektronenkonzentration der Anreicherungsschicht gerechnet werden. Es werden deshalb zur rechnerischen Vereinfachung im folgenden die spannungsabhängigen mittleren Trägerkonzentrationen in der Anreicherungsschicht, die sich aus einer elektrostatischen Belastung der Inversionsschicht senkrecht zur Oberfläche berechnen, auch für die strombelastete Inversionsschicht bei Existenz eines Oberflächenstromes zugrunde gelegt.

Bisher wurden nur Fälle betrachtet, bei denen die Struktur der Oberflächeninversionsschicht homogen über die ganze Oberfläche verändert wird. Für die spätere Betrachtung ist jedoch die Einwirkung eines an jeder Stelle der Oberfläche verschiedenen äußeren Potentials von Bedeutung.

Dieser Fall der Einwirkung eines äußeren inhomogenen elektrischen Feldes ist ebenfalls schematisch in Bild 7 dargestellt. Bild 7c zeigt ein Beispiel rein elektrostatischer Belastung, realisiert durch schräge Annäherung einer geladenen Elektrode an die Halbleiteroberfläche. In Bild 7d ist das Schema eines dünnen länglichen Halbleiterplättchens mit einer strombelasteten Oberflächeninversionsschicht wiedergegeben (Fall eines stationären Nichtgleichgewichtszustandes). In beiden Fällen werden als Demonstrationsbeispiel die Verhältnisse bei Flußbelastung skizziert. Man sieht die gegensinnige Tendenz im Verlauf der Grenzen der randnahen Anreicherungsschicht (schraffiert gezeichnet) und der Raumladungszone (gestrichelte Linie). Im Fall

von Bild 7d ist das n-leitende Halbleiterplättchen kontaktiert und in einen äußeren Stromkreis eingeschlossen. Die Einwirkung eines inhomogenen äußeren Potentials auf die Oberflächeninversionsschicht kommt hier dadurch zustande, daß die Oberflächeninversionsschicht AB bei A an einen die Kristalloberfläche an der Stirnseite des Plättchens durchsetzenden konventionellen p-n-Übergang sperrfrei anschließt und somit parallel zu diesem Volumen-p-n-Übergang geschaltet ist. Der Potentialabfall zwischen den Punkten A und B längs der Oberflächenschicht, der gleich dem Potentialabfall über dem gewöhnlichen p-n-Übergang an der Stirnseite des Plättchens ist, ruft an jeder Stelle der Oberfläche in der Randschicht eine verschiedene Potentialdifferenz in x -Richtung hervor, da der gesamte n-leitende Bahnbereich des Kristalls bei kleinen Belastungen des Volumen-p-n-Überganges praktisch auf dem gleichen Potential wie der Punkt B liegt. Die aufgeprägte Belastung der Oberflächeninversionsschicht in x -Richtung ist somit an jeder Stelle der Halbleiteroberfläche verschieden. Eine solche Anordnung läßt sich auch experimentell realisieren.

Für die physikalischen Verhältnisse des Oberflächenstromes in der Anreicherungsschicht parallel zur Halbleiteroberfläche sind insbesondere die folgenden Umstände von Bedeutung. Infolge der inhomogenen äußeren Belastung ändert sich der Konzentrationsverlauf der Defektelektronen in der Anreicherungsschicht an jeder Stelle der Halbleiteroberfläche verschieden. Dadurch ist die Dicke der Anreicherungsschicht, ihre spannungsabhängige Trägerkonzentration, gemittelt senkrecht zur Oberfläche, und der durch die Belastung hervorgerufene Strom senkrecht zur Oberfläche an jeder Stelle der Oberfläche verschieden groß. (Der Verlauf der senkrecht zur Halbleiteroberfläche gemittelten Defektelektronenkonzentration der Anreicherungsschicht parallel zur Oberfläche ist bei relativ kleinen Belastungen im elektrostatischen Fall (Bild 7c) und im strombehafteten Fall (Bild 7d) praktisch der gleiche.)

Aus diesem Grunde ist es zweckmäßig, für eine quantitative Beschreibung des Oberflächenstromes bei starker Inversion die Konzeption einer vereinfachten Modellinversionsschicht zu verwenden. Der experimentell gemessene Oberflächenstrom läßt sich in einer analytisch geschlossenen Form in diesem Bild beschreiben, wobei die Modellanreicherungsschicht an jeder Stelle der Oberfläche als effektive Trägerdichte die dort existierende spannungsabhängige mittlere invertierte Teilchenkonzentration (gemittelt senkrecht zur Oberfläche) enthält. Gerechtfertigt wird diese Methode dadurch, daß experimentell im wesentlichen diese mittlere invertierte Trägerkonzentration an jeder Stelle der Anreicherungsschicht als effektive Größe den Oberflächenstrom quantitativ bestimmt.

Der nichtlineare Potentialverlauf zwischen A und B parallel zur Halbleiteroberfläche ist kausal verknüpft mit einem äquivalenten Verlauf der mittleren, invertierten Trägerdichte der Modellanreiche-

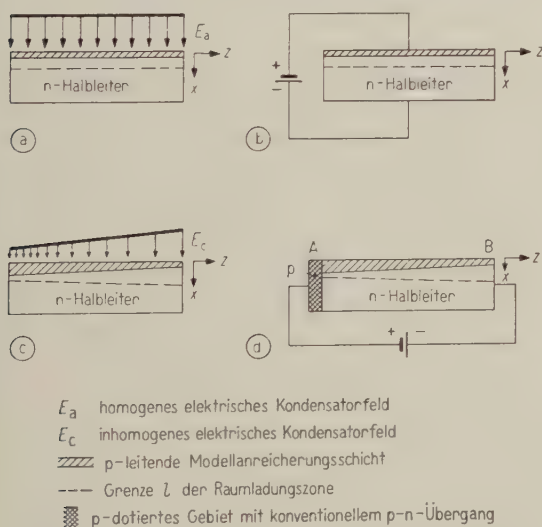


Bild 7. Schematische Darstellung der Beeinflussung von Oberflächen-Inversionsschichten durch äußere Felder (Flußrichtung);

- homogene elektrostatische äußere Gleichbelastung,
- homogener Nicht-Gleichgewichtszustand einer idealisierten, strombehafteten Gleichbelastung,
- inhomogene elektrostatische äußere Gleichbelastung,
- inhomogener Nicht-Gleichgewichtszustand einer strombehafteten Gleichbelastung.

Bild 8. Modell einer Spitzendiode mit Oberflächen-Inversionsschicht.

der Spannung als nicht abhängig angenommenen halbkugelförmigen p-Zone des Formiergebietes ein sperrfreier Übergang ohne zusätzliche Bedingungen angenommen. An der äußeren sperrfrei an den Halbleiter kontaktierten Elektrode r_2 wird die an sich willkürliche äußere Spannung Null festgehalten, so daß hier die Verhältnisse in der Anreicherungsschicht nicht durch eine an den Spitzenkontakt angelegte Spannung modifiziert werden.

Sowohl im Sperrfall als auch bei kleinen Flußbelastungen, wie sie hier experimentell verwirklicht wurden, erfolgt der gesamte Spannungsabfall des Volumenstromes an der halbkugelförmigen Sperrschicht. Demzufolge herrscht im gesamten n-leitenden Bahngebiet ein konstantes Potential von gleicher Größe wie an der Außenelektrode r_2 . In der Anreicherungsschicht hingegen existiert wegen des ohmschen Überganges an der Stelle r_1 bei Belastung der Diode eine Potentialverteilung parallel zur Oberfläche bis zum Rand r_2 . Unter den obigen Voraussetzungen ist somit an jeder Stelle r die jeweilige Potentialdifferenz zwischen r und r_2 ebenso groß wie das an dieser Stelle r senkrecht zur Halbleiteroberfläche an der gesamten Raumladungszone der Oberflächeninversionsschicht liegende äußere Potential.

Da dieses von r abhängige äußere Potential an jeder Stelle r mit verschiedener Größe den Aufbau der gesamten Raumladungszone der Oberflächeninversionsschicht senkrecht zur Oberfläche sowie die Dicke der Anreicherungsschicht und die Konzentrationsverteilung ändert, ist das Schema der unbelasteten Modellanreicherungsschicht des Bildes 8 somit bei Belastung durch eine annähernd keilförmige Schicht ersetzt zu denken.

4. Der Mechanismus des Oberflächenstromes

Der experimentell erfaßbare Oberflächenstrom wird bei starker Inversion im wesentlichen durch die spannungsabhängige mittlere Defektelektronenkonzentration der Anreicherungsschicht (gemittelt jeweils senkrecht zur Oberfläche) repräsentiert. Wir betrachten deshalb, wie bereits oben begründet, eine Modellanreicherungsschicht, die im unbelasteten Gleichgewichtszustand eine homogene, über den wirklichen Konzentrationsverlauf senkrecht zur Oberfläche gemittelte, effektive Trägerkonzentration besitzt. Der Mechanismus des Oberflächenstromes im Bild des vereinfachten Modells beruht nun darauf, daß sich beim Anlegen einer äußeren Spannung an die Diode in der Anreicherungsschicht parallel zur Oberfläche eine zylindersymmetrische Potentialverteilung ausbildet, wobei das äußere Potential an jeder Stelle r eine bestimmte Änderung der effektiven mittleren Defektelektronendichte hervorbringt.

Somit entspricht der Potentialverteilung wechselseitig auch eine Verteilung der Variation der mittleren Trägerdichte. In der Modellanreicherungsschicht wird für die durch die zylindersymmetrische Potentialverteilung längs der Oberfläche verursachten Trägerbewegungen die vereinfachende Voraus-

setzung gemacht, daß die Variation $\delta\bar{p}$ der mittleren Defektelektronenkonzentration gleich der Variation $\delta\bar{n}$ der mittleren Elektronenkonzentration ist. Diese Annahme erscheint zunächst hinsichtlich der Raumladungsverhältnisse in der Inversionsschicht, senkrecht zur Oberfläche betrachtet, als eine drastische Vereinfachung. Man muß jedoch hierbei berücksichtigen, daß die sich auf den Aufbau der Inversionsschicht senkrecht zur Oberfläche beziehenden analytischen Lösungen von dieser Voraussetzung nicht betroffen werden [17]. Ist aber die senkrecht zur Oberfläche gemittelte Trägerkonzentration $\delta\bar{p}$ als Funktion der jeweils existierenden äußeren Spannung ermittelt, so läßt die zur Lösung des Randwertproblems der Trägerverteilung $\delta\bar{p}(r)$ längs der Oberfläche gemachte, mathematisch vereinfachende Annahme $\delta\bar{p} = \delta\bar{n}$ sich in der Modellanreicherungsschicht gemäß ihrer Definition auch physikalisch im Sinne einer zweckmäßigen Näherung rechtfertigen [19], da sich die Modellanreicherungsschicht in bezug auf den Verlauf der Trägervariation $\delta\bar{p}(r)$ längs der Oberfläche analog verhält wie ein trägermoduliertes quasineutrales Bahngebiet bei hoher Injektion. Wo keine Trägervariation stattfindet, herrscht auch keine äußere elektrische Kraft und umgekehrt. Die Wirkung des äußeren Potentials U auf die Anreicherungsschicht kommt somit einerseits in der Änderung des inneren Potentials der Raumladungszone in Richtung senkrecht zur Oberfläche mit einer entsprechenden Variation der mittleren Defektelektronenkonzentration ($-\partial U/\partial\delta\bar{p}$) und andererseits in einem hierdurch bedingten Diffusionsgradienten ($\partial\delta\bar{p}/\partial r$) parallel zur Oberfläche zum Ausdruck. Beide Wirkungen zusammen ergeben in der Modellanreicherungsschicht parallel zur Oberfläche einen Feld- und einen Diffusionsstromanteil. Der Feldstromanteil, der durch die Variation der mittleren Trägerdichte wechselseitig bedingt ist, wird an jeder Stelle r durch die Teilnahme der jeweils bei r vorhandenen gesamten beweglichen Träger der Anreicherungsschicht gebildet. Am Diffusionsstrom nehmen nur die Ladungsträger teil, die die Variation $\delta\bar{p}$ bzw. $\delta\bar{n}$ verursachen. Der Diffusionsgradient, der diesen Strom bestimmt, hängt von der Größe der Trägervariation und deren zylindersymmetrischer Verteilung parallel zur Oberfläche ab.

Für die experimentelle Größe des Oberflächenstromes der Spitzendiode ist folgende Überlegung maßgebend: Der bei r_1 in der Modellanreicherungsschicht parallel zur Oberfläche gerichtete Strom ist identisch mit dem in Wirklichkeit in komplizierten Stromlinien zwischen r_1 und r_2 über die gesamte Oberflächeninversionsschicht in das n-leitende Halbleiterinnere fließenden Oberflächenstrom. Bei r_2 verschwindet sowohl die Trägervariation als auch das elektrische Feld parallel zur Oberfläche, und es fließt somit kein Strom an dieser Stelle. Auf Grund dieser Randbedingungen ist nur der durch die Potentialänderung erzeugte Diffusionsgradient parallel zur Oberfläche und die Variation der Träger, beide Größen am Rande r_1 , zur Berechnung des Ober-

flächenstromes notwendig. Die physikalischen Verhältnisse des Diodenmodells von Bild 8 sind vollkommen analog zu dem oben erläuterten Beispiel der Halbleiterstruktur in Bild 7d.

5. Die Strom-Spannungskennlinie des Oberflächenstromes

Die physikalischen Ursachen der Kennliniengestalt und der Mechanismus des Oberflächenstromes sollen an Hand von Bild 9 anschaulich gemacht und erläutert werden.

Zur quantitativen Berechnung des Oberflächenstromes genügt die Berechnung des Stromes, der an der Stelle r_1 in die Anreicherungsschicht hineinfließt. Als Beispiel wurden die folgenden Werte für die Halbleiterparameter gewählt: n-Germanium mit einer Dotierung $n_D = 10^{15}$ Atome/cm³ bei 300°K, Randfeldstärke an der Halbleiteroberfläche $E_R = 10^4$ V/cm, das Verhältnis zwischen Radius r_1 des formierten Halbkugelgebietes und der effektiven Diffusionslänge L_{ob} der Anreicherungsschicht $r_1/L_{ob} = 0,125$. Eine Modifikation der Beweglichkeit in der Anreicherungsschicht gegenüber dem Wert im Halbleiterinneren wurde in diesem Beispiel nicht in Betracht gezogen.

In Bild 9 sind in Abhängigkeit von der Spannung die Größen aufgetragen, die gemäß der im Anhang angegebenen Formel (6) die Gestalt der Strom-Spannungs-Kennlinie bestimmen. Der berechnete funktionale Zusammenhang bezieht sich auf die

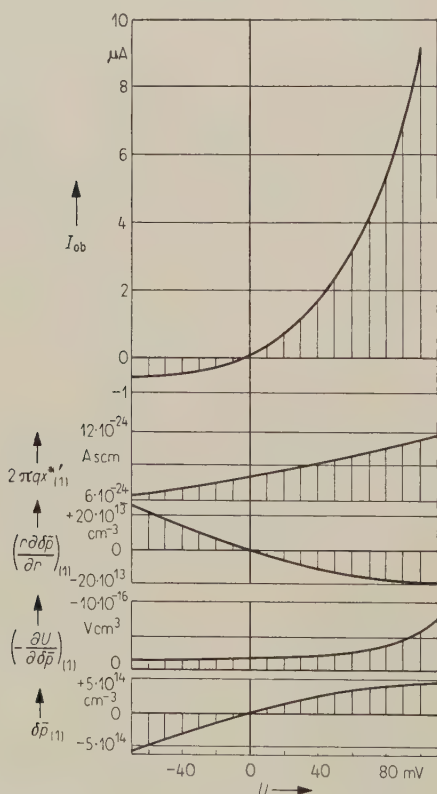


Bild 9. Zur Kennliniengestalt des Oberflächenstromes; n-Germanium, $T = 300^\circ\text{K}$, $n_D = 10^{15}$ Atome/cm³, $E_R = 10^4$ V/cm, $r_1/L_{ob} = 0,125$.

Verhältnisse bei kleinen Gleichspannungen. Zunächst ist der Kurvenverlauf der Variation $\partial\bar{p} = \partial n$ der mittleren Defektelektronenkonzentration an der Stelle $r = r_1$ der Anreicherungsschicht dargestellt. Die mittlere Trägerdichte $\partial\bar{p}$ ändert sich mit der angelegten Spannung in Fluß- und Sperrrichtung unsymmetrisch. Das Absinken der mittleren Trägerkonzentration bei Sperrbelastung ist durch die Anzahl der überhaupt in der Anreicherungsschicht vorhandenen Teilchen grundsätzlich begrenzt. Bei hinreichend großer Sperrspannung kann die Defektelektronenkonzentration in der Anreicherungsschicht soweit absinken, daß diese praktisch verschwindet. Der Oberflächenkanal wird dann abgeschnürt. In Flußrichtung besteht eine solche durch die Anzahl der Teilchen bedingte Grenze im Prinzip nicht. Jedoch steigt die mittlere Trägerkonzentration mit größer werdender Spannung, wie aus dem Diagramm zu ersehen ist, nur noch langsam. Es stellt sich schließlich eine Art Sättigungszustand ein, da das Anwachsen der mittleren Defektelektronenkonzentration mit steigender Flußspannung durch die sich einstellende Raumladungsverteilung begrenzt wird. Durch das Verhalten der Variation $\partial\bar{p}$ ist auch der Verlauf ihres negativen Differentialquotienten bzw. dessen reziproken Wertes in Abhängigkeit von der Spannung mitbestimmt. Bemerkenswert ist die relativ starke Änderung dieser Ableitung im Flußgebiet.

Der spannungsabhängige Ausdruck $(r\partial\delta\bar{p}/\partial r)_{(1)}$ repräsentiert den Diffusionsgradienten der mittleren Trägerdichte am Rande $r = r_1$, multipliziert mit r_1 . Bei Flußspannung ist der Diffusionsgradient bei $r = r_1$ negativ, weil die Abweichung $\partial\bar{p}$ der mittleren Trägerkonzentration vom Gleichgewichtswert \bar{p}_0 dort ihren größten Wert annimmt und mit zunehmendem r allmählich bis auf $\partial\bar{p} = 0$ bei $r = r_2$ abklingt. Eine Sperrspannung bewirkt eine Abnahme der mittleren Defektelektronenkonzentration, so daß der Diffusionsgradient bei $r = r_1$ positiv wird. Schließlich ist in Bild 9 noch die Dicke x^* (Stern bedeutet Spannungsabhängigkeit) der Anreicherungsschicht an der Stelle $r = r_1$, multipliziert mit dem Faktor $2\pi q$ (q Betrag der Elementarladung), in Abhängigkeit von der Spannung aufgetragen. Aus dem Verlauf der verschiedenen Größen ergibt sich dann gemäß Formel (6) die dargestellte Strom-Spannungs-Charakteristik des Oberflächenstromes [19].

6. Experimentelle Untersuchung

Die Kennlinienmessungen wurden mit einer Kompensationsmethode durchgeführt, die einerseits eine genaue Bestimmung von Meßwerten auch bei kleinen Strömen erlaubt und andererseits Fehler durch Innenwiderstände der Meßinstrumente zu eliminieren gestattet. Für die Dioden wurden Germanium-Plättchen verwendet, die mit Arsen dotiert waren. Als Ausgangsmaterial diente ein aus der Schmelze gezogener Germanium-Einkristallstab. Die Störstellendichte betrug $n_D = 0,8 \cdot 10^{15}$ cm⁻³, was einem spezifischen Widerstand von $2 \Omega\text{cm}$ ent-

spricht. Als Federmaterial wurde Molybdän gewählt. Der Krümmungsradius der Molybdänspitze schwankte bei den einzelnen Dioden zwischen $3\text{ }\mu\text{m}$ und $5\text{ }\mu\text{m}$. Bei der Formierung wurde die Federkraft der Metallschmelze mit Hilfe eines Dynamometers für alle Dioden gleich groß eingestellt. Der Formierstromimpuls betrug $0,6\text{ A}$ und $0,75\text{ s}$ und war für sämtliche Exemplare gleich.

Um die gemessenen Kurvenverläufe mit den theoretischen Kurven vergleichen zu können, ist die Kenntnis einer Reihe von Parametern erforderlich: Der Radius r_1 der Formierhalbkugel, die Diffusionslänge im Kristall L_{vol} und in der Anreicherungsschicht der Oberflächeninversionsschicht L_{ob} , die effektive elektrische Flächenladungsdichte auf der Kristalloberfläche bzw. die äquivalente Randfeldstärke. Der Radius r_1 variiert mit den Formierungsbedingungen und repräsentiert somit das Ergebnis einer bestimmten Formiervorschrift. Die Diffusionslänge L_{vol} charakterisiert den Leitungsmechanismus im Inneren des Kristalls und bestimmt somit den über die Formierhalbkugel fließenden Volumenanteil des Diodenstromes. Die Größe L_{vol} wird durch Messungen am Ausgangskristall bestimmt, der zur Herstellung der Dioden verwendet wurde. Durch die elektrische Randfeldstärke ist bei vorgegebener Dotierung der Aufbau der Oberflächeninversionsschicht festgelegt. Die elektrische Flächenladungsdichte an der Oberfläche und somit die Randfeldstärke kann durch Oberflächenbehandlung bei der Herstellung der Spitzendioden verändert werden. Für die fertige Diode stellt die elektrische Flächenladungsdichte eine charakteristische effektive Größe dar, durch welche die physikalische Beschaffenheit der Kristalloberfläche und ihre Wirkung auf den Gleichrichtereffekt zum Ausdruck kommt. Desgleichen ist die Diffusionslänge L_{ob} in der an den Kristallrand anschließenden Anreicherungsschicht als eine effektive phänomenologische Kenngröße des Oberflächenstromes der Spitzendiode anzusehen.

7. Analyse experimenteller Kurven

An Hand theoretisch ermittelter Strom-Spannungs-Kennlinien sollen jetzt experimentelle Kurven untersucht werden, die an Spitzendioden gemessen wurden. Dabei standen die experimentellen Voraussetzungen im Einklang mit den theoretischen Grundannahmen der im Anhang zusammengestellten analytischen Ergebnisse.

In Bild 10 sind zunächst zwei gemessene Strom-Spannungs-Kennlinien durch Meßpunkte dargestellt. Die berechnete Kurve für den Gesamtstrom I_{ges} ist als ausgezogene Linie aufgetragen. Hierzu wurden folgende Modellannahmen im Einklang mit den experimentellen bekannten Daten gemacht: Dotierung $n_D = 0,8 \cdot 10^{15}\text{ cm}^{-3}$, Diffusionslänge für den Volumenstromanteil $L_{\text{vol}} = 400\text{ }\mu\text{m}$, äußerer Radius des n-leitenden Kristalls $r_2 = 800\text{ }\mu\text{m}$. Außerdem wurden zur Anpassung der theoretischen Kurve an die gemessenen die folgenden Bestimmungsgrößen für das Modell gewählt: Randfeld-

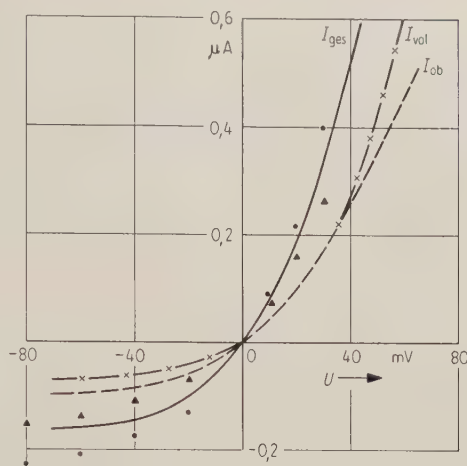


Bild 10. Berechneter Volumen-, Oberflächen- und Gesamtstrom in Abhängigkeit von der angelegten Spannung im Vergleich zu gemessenen Spitzengleichrichter-Kennlinien;

n-Germanium, $T = 300^\circ\text{K}$, $n_D = 0,8 \cdot 10^{15}\text{ Atome pro cm}^3$, $\rho = 2\text{ }\Omega\text{ cm}$, $E_R = 0,8 \cdot 10^4\text{ V/cm}$, $L_{\text{ob}} = 320\text{ }\mu\text{m}$, $L_{\text{vol}} = 400\text{ }\mu\text{m}$, $r_1 = 20\text{ }\mu\text{m}$, $r_2 = 800\text{ }\mu\text{m}$, $D_{\text{pob}} = 15\text{ cm}^2/\text{s}$, $D_{\text{nob}} = 30\text{ cm}^2/\text{s}$, $\mu_{\text{ob}} = \mu_{\text{vol}}/3$; gemessene Werte: \blacktriangle Diode Nr. 1, \bullet Diode Nr. 2.

stärke $E_R = 0,8 \cdot 10^4\text{ V/cm}$, diesem Wert entspricht bei Germanium eine elektrische Flächenladungsdichte an der Kristalloberfläche von etwa 10^{11} Elementarladungen pro cm^2 , was größenordnungsmäßig auch von anderen Autoren ermittelt wurde [20]. Für die Anpassung der theoretischen Kennlinie an die gemessene wurde als effektive Diffusionslänge des Oberflächenstromes in der Anreicherungsschicht der Wert $L_{\text{ob}} = 320\text{ }\mu\text{m}$ benutzt; dieser Wert ist kleiner als im Inneren des Kristalls, um den randnahen Verhältnissen Rechnung zu tragen. Die Beweglichkeit der Ladungsträger in der Anreicherungsschicht wurde entsprechend den SCHRIEFFERschen Untersuchungen gegenüber dem Inneren um $2/3$ verringert [21]. Als Radius der Formierhalbkugel wurde $r_1 = 20\text{ }\mu\text{m}$ angenommen. Dies stimmt mit den Ergebnissen von THEDECK [22] überein und ist außerdem durch unveröffentlichtes Erfahrungsmaterial in befriedigender Weise belegt. Außer der theoretischen Kurve für den Gesamtstrom sind in Bild 10 die berechneten Kennlinien des zugehörigen Volumen- und Oberflächenstromes dargestellt.

Wie aus der graphischen Darstellung hervorgeht, ergibt sich im Durchlaßbereich, soweit die Messung durchgeführt wurde, und im Sperrbereich bis etwa 50 mV eine befriedigende Übereinstimmung der theoretischen mit den gemessenen Kurven. Für größere Sperrspannungen weichen die theoretischen und experimentellen Kurven etwas voneinander ab. Der gemessene Strom nimmt dann stärker zu als der theoretisch ermittelte. BENEKING [23] hat nachgewiesen, daß bei Berücksichtigung der Spannungsabhängigkeit der Sperrschichtdicke des halbkugelförmigen p-n-Überganges der Sperrstrom mit wachsender Spannung steigt. In den der vorliegenden Arbeit zugrunde liegenden Berechnungen ist dieser Effekt des Volumenstromes aus Gründen der Ver-

einfachung des Modells nicht berücksichtigt worden. Es liegt jedoch nahe, für den zunehmenden Anstieg des Sperrstromes noch eine Reihe anderer physikalischer Gründe in Betracht zu ziehen. Beispielsweise ist es denkbar, daß die zunehmende äußere Sperrspannung die wirklichen Bedingungen an der Kristalloberfläche stärker modifiziert, als es den verwendeten theoretischen Voraussetzungen des Modells entspricht, und dadurch den Oberflächenstrom beeinflusst.

Es soll noch bemerkt werden, daß zur Angleichung der berechneten an die gemessenen Kurven bereits ein relativ kleiner Radius $r_1 = 20 \mu\text{m}$ des formierten Halbkugelgebietes ausreicht. Würde man den gemessenen Gesamtstrom allein durch den Volumenstrom beschreiben, so müßten hierbei Werte r_1 von etwa $50 \mu\text{m}$ bis $100 \mu\text{m}$ verwendet werden, was nicht in Einklang mit experimentellen Gegebenheiten steht.

8. Vergleich der theoretischen Ergebnisse mit einer gemessenen Kennlinienschar

Bei dem Vergleich in Bild 10 zwischen theoretischen und experimentellen Ergebnissen sind zur Demonstration des Oberflächeneinflusses nur zwei gemessene Diodenkennlinien betrachtet worden. Jetzt soll eine gemessene Kennlinienschar zur Analyse herangezogen werden. Die hierbei verwendeten Spitzenkontakt-Gleichrichter sind alle aus dem gleichen Kristall mit einer Donatorenkonzentration von $0,8 \cdot 10^{15} \text{ Atome/cm}^3$ hergestellt ($L_{\text{vol}} = 400 \mu\text{m}$), und sie wurden bis zur Messung in jeder Hinsicht streng den gleichen äußeren Bedingungen ausgesetzt.

Die Daten für den bei allen Exemplaren gleichen Formierprozeß sind bereits oben mitgeteilt worden. Bemerkenswert ist, daß trotz dieser sorgfältig beachteten Gleichheit der äußeren Bedingungen keine der gemessenen Kennlinien in Strenge einer anderen gleicht.

Dieser Sachverhalt demonstriert ein bekanntes Phänomen, das für die Produktion von Halbleiterbauelementen von großer Bedeutung ist. Sowohl die phänomenologischen Zustandsgrößen, die den Herstellungsprozeß des Bauelements kennzeichnen, als auch die physikalischen Bestimmungsgrößen des fertigen Exemplars reichen nicht aus, um die elektrischen Eigenschaften eindeutig zu definieren. Jeder Kristall und insbesondere jedes Halbleiterbauelement besitzt eine gewisse Individualität, die für differenziertere Zwecke erst genau bestimmt werden muß und die vor allem durch Oberflächeneinflüsse hervorgerufen wird. Diese Einflüsse werden sich bei einem Exemplar um so stärker bemerkbar machen, je mehr der Mechanismus des Oberflächenstromes an der gesamten Strombilanz beteiligt ist. Bei Spitzenkontakt-Gleichrichtern spielt im Vergleich zu anderen Arten von Bauelementen, wie eingangs schon ausgeführt, infolge der geometrischen Anordnung der Oberflächenstrom eine besonders entscheidende Rolle. Die Analyse der Kennlinien durch Vergleich mit berechneten Strom-Spannungs-Kurven läßt eine planmäßige Beein-

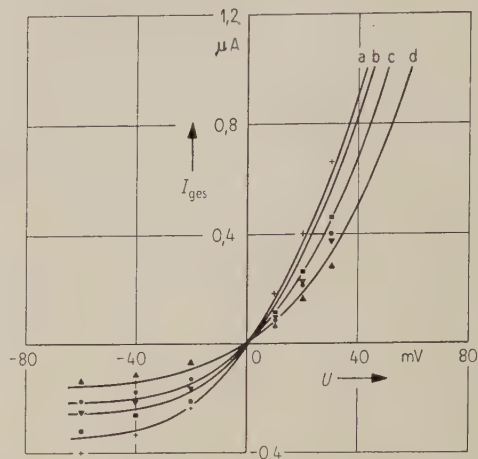


Bild 11. Theoretische Kennlinien mit verschiedenen Parametern im Vergleich zu einem durch Meßpunkte dargestellten experimentellen Kennlinienfeld; n-Germanium, $n_D = 0,8 \cdot 10^{15} \text{ Atome/cm}^3$, $T = 300^\circ\text{K}$, $\rho = 2 \Omega \text{ cm}$.

Kurve	E_R	L_{ob}	L_{vol}	r_1	r_2
	V/cm	μm	μm	μm	μm
a	10^4	300	400	12	800
b	$0,8 \cdot 10^4$	320	400	25	800
c	$0,8 \cdot 10^4$	320	400	20	800
d	$0,8 \cdot 10^4$	320	400	20	800

Kurve d außerdem: $D_{\text{pob}} = 15 \text{ cm}^2/\text{s}$,
 $D_{\text{nob}} = 30 \text{ cm}^2/\text{s}$,
 $\mu_{\text{ob}} = \mu_{\text{vol}}/3$.

Gemessene Werte: \blacktriangle Diode Nr. 1, \bullet Diode Nr. 2,
 \blacksquare Diode Nr. 3, \blacktriangledown Diode Nr. 4,
 $+$ Diode Nr. 5.

flussung der Kennlinienform durch entsprechende Wahl der wesentlichen Bestimmungsgrößen in den Bereich der Möglichkeiten treten.

In Bild 11 sind in das durch Meßpunkte angezeigte Kennlinienfeld vier theoretische Strom-Spannungs-Kennlinien des Gesamtstromes eingetragen. Daß hier nur fünf Kennlinien aufgezeichnet sind, hat nur den Grund besserer Übersichtlichkeit. Die gewählten Kurven repräsentieren eine größere Anzahl ähnlicher Kurven innerhalb des gleichen Streubereiches. Der Vergleich zwischen den theoretischen und gemessenen Kurven zeigt, in welcher Weise die Streuung der gemessenen Kennlinien durch Variation der verfügbaren Parameter beispielsweise zustandekommen kann. Bei den gemessenen Kurven sind folgende Parameter bekannt: n-Germanium mit einer Dotierung $n_D = 0,8 \cdot 10^{15} \text{ Atome/cm}^3$ entsprechend $\rho = 2 \Omega \text{ cm}$, $T = 300^\circ\text{K}$, $L_{\text{vol}} = 400 \mu\text{m}$, $r_2 = 800 \mu\text{m}$. Diese Werte wurden auch in den theoretischen Kennlinien zugrunde gelegt. Für die noch verfügbaren Parameter wurden die in der Unterschrift zu Bild 11 angegebenen, im Bereich der Erfahrung liegenden Annahmen gemacht.

Bei Kurve d wurde die Beweglichkeit der Ladungsträger in der Anreicherungsschicht der Oberfläche gegenüber a, b und c um $2/3$ des konventio-

nellen Wertes bei Germanium vermindert. In den Kurven a, b und c wurde für die Beweglichkeit in der Anreicherungsschicht näherungsweise der konventionelle Wert des Halbleiterinneren verwendet. Eine Modifikation der Beweglichkeit bei gleichbleibender Diffusionslänge entspricht einer Änderung der mittleren Lebensdauer der Ladungsträger.

In den vorliegenden Beispielen sind nur Oberflächenströme betrachtet worden, die sich auf Grund starker Inversion [17] ergeben. Entsprechende Betrachtungen lassen sich aber auch für den Fall schwacher Inversion [17] durchführen. Es ist sogar wahrscheinlich, daß bei Spitzendioden, die unter gleichen äußeren Bedingungen hergestellt wurden, beide Fälle und dazwischenliegende Übergänge vorkommen.

9. Allgemeine Gesetzmäßigkeiten

Hinsichtlich des Einflusses der physikalischen Parameter auf die Kennlinien lassen sich allgemein folgende Gesetzmäßigkeiten erkennen. Der Oberflächenstrom und damit auch der Gesamtstrom wächst im Fluß- und Sperrgebiet, wenn die Randfeldstärke E_R bzw. die elektrische Flächenladungsdichte im unbelasteten Gleichgewichtszustand an der Oberfläche erhöht oder der Radius r_1 vergrößert oder die Oberflächendiffusionslänge in der Anreicherungsschicht L_{ob} verringert wird. Die übrigen Parameter sind hierbei festzuhalten. Eine Vergrößerung des Radius r_1 bedeutet gleichzeitig eine Zunahme des Volumenstromanteils. Eine Verringerung der Beweglichkeit in der Anreicherungsschicht der Oberfläche hat außerdem eine Verkleinerung der Oberflächenleitfähigkeit und somit auch des Oberflächenstromes zur Folge. Man sieht hieraus, wie sich der Einfluß der Oberflächenverhältnisse auf das gesamte elektrische Verhalten der Spitzkontakt-Gleichrichter bemerkbar macht.

Aus den dargestellten Kurvenverläufen ist ersichtlich, daß sowohl im Sperr- als auch im Durchlaßbereich der Oberflächenstrom den größten Anteil am Gesamtstrom liefern kann. Wegen der grundsätzlichen Bedeutung der Oberflächeneinflüsse auf den Diodenstrom ist in Bild 12 der gesamte Sperrstrom in Abhängigkeit von der Diffusionslänge L_{ob} der Anreicherungsschicht aufgetragen. Physikalisch entspricht einer größeren Diffusionslänge eine geringere Rekombination in der Modellanreicherungsschicht. Die Rekombinationsfähigkeit der Oberfläche kann durch verschiedene Oberflächenbehandlung beeinflusst werden. Die Kurve zeigt den Verlauf bei einer willkürlich herausgegriffenen Sperrspannung von $U = -40$ mV. Mit zunehmender Diffusionslänge des Oberflächenstromes wird der Betrag des Sperrstromes kleiner. Dieser Sachverhalt steht in Analogie zur konventionellen Theorie des p-n-Überganges. Von praktischer physikalischer Bedeutung ist jedoch, daß sich der Sperrstrom außerordentlich vergrößert, wenn die effektive Diffusionslänge in der Anreicherungsschicht wesentlich kleiner als im Kristallinneren wird. Es läßt sich also bei sonst gleichbleibenden Bedingungen ein

möglichst kleiner Sperrstrom erzielen, indem man durch entsprechende Oberflächenbehandlung die Rekombinationsfähigkeit der Oberfläche wirksam herabsetzt und hierdurch die Diffusionslänge L_{ob} hinreichend groß macht.

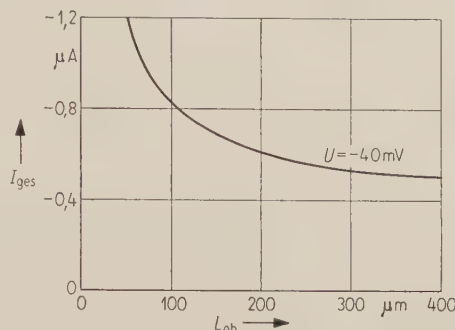


Bild 12. Berechneter Gesamtdiodenstrom in Abhängigkeit von der effektiven Diffusionslänge der Anreicherungsschicht an der Halbleiteroberfläche;

n-Germanium, $n_D = 10^{15}$ Atome/cm³, $\rho = 1,5 \Omega \text{ cm}$, $T = 300^\circ \text{ K}$, $E_R = 10^4 \text{ V/cm}$, $L_{vol} = 400 \mu\text{m}$, $= \text{const}$, $r_1 = 25 \mu\text{m}$, $r_2 = 800 \mu\text{m}$.

10. Zur Frage des Richtverhältnisses

Wird die Kennlinie eines Spitzkontakt-Gleichrichters ohne Berücksichtigung des Oberflächenstromes nur durch den über das formierte Halbkugelgebiet fließenden Volumenstrom erzeugt, so ist das von der Spannung U abhängige statische Richtverhältnis r im Rahmen der SHOCKLEYSchen Theorie durch den Parameter $\alpha = q/kT$ eindeutig bestimmt:

$$r_{vol} = \frac{I_{d vol}}{I_{sp vol}} = e^{\alpha U}. \quad (1)$$

Im Anhang sind in den Formeln (2) bis (5) die hier einzusetzenden analytischen Ausdrücke für den Volumen-Durchlaßstrom $I_{d vol}$ und Volumen-Sperrstrom $I_{sp vol}$ angegeben. Das Richtverhältnis variiert demzufolge bei vorgegebener Spannung nur mit der Temperatur. Dieser Sachverhalt ist in Bild 13 dargestellt. In Abhängigkeit von der Spannung U ist hier das Richtverhältnis bei verschiedenen α -Werten aufgetragen. Die den Parameterwerten α zugehörigen Temperaturen sind in dem Diagramm mitangegeben. Mit abnehmendem α wächst die Temperatur. Bei einer Zimmertemperatur von 27° C hat α den Wert $38,6 \text{ V}^{-1}$. Die zugehörige Richtverhältniskurve hat in diesem Fall, bei dem die Temperatur der Diode mit der Umgebungstemperatur übereinstimmt, unter gewöhnlichen experimentellen Umständen den größten α -Wert und demzufolge die günstigsten Richtverhältniswerte. Diese der Zimmertemperatur entsprechende Kurve begrenzt die Kurvenschar des Bildes 13 nach oben. Ermittelt man aus experimentellen Kennlinien von Spitzkontakt-Gleichrichtern die zugehörigen Kurven des Richtverhältnisses und trägt diese in die Kurvenschar von Bild 13 ein, so zeigt sich, daß die gemessenen Kennlinien α -Werte aufweisen, die zum Teil beträchtlich kleiner als 38,6 sind. Den experi-

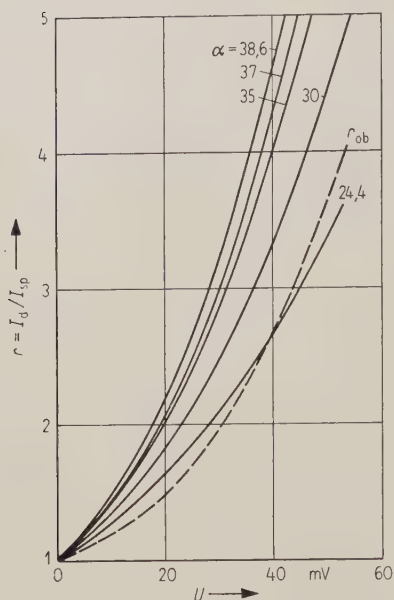


Bild 13. Theoretische Richtverhältnisse als Funktion der Spannung bei verschiedenen Temperaturen. Ausgezogene Kurven für Volumenstrom, gestrichelte Kurve für Oberflächenstrom.

α	38,6	37	35	30	24,4	V ⁻¹
t	27	40	58	113	202	°C

mentellen Kennlinien müßten demzufolge Diodentemperaturen zugeordnet werden, die im allgemeinen wesentlich höher als die in Wirklichkeit existierende Temperatur sind.

Bei den in der vorliegenden Arbeit verwendeten experimentellen Kurven sind die aus den abgelesenen α -Werten entnommenen relativ hohen Temperaturen mit Sicherheit nicht mit den vorherrschenden Verhältnissen geringer Belastungen zu vereinbaren. Neben anderen Deutungen erklärte H. A. BETHE die kleinen experimentellen α -Werte durch eine Superposition von Kennlinien einzelner Gleichrichterelemente [24]. Die den gemessenen Kurven zugeordneten niedrigen α -Werte bzw. die entsprechenden Übertemperaturen werden offensichtlich durch einen zusätzlichen Effekt vorgetauscht, der im Rahmen der für den Volumenstrom gültigen Theorie des p-n-Überganges nicht explizit enthalten ist.

Es läßt sich zeigen, daß sich bei Berücksichtigung des Oberflächenstromes diese Schwierigkeiten grundsätzlich nicht ergeben. Der in Gl. (1) nicht erfaßte Effekt kann durch den Oberflächenstrom eine befriedigende Erklärung finden. In Bild 13 ist in die Kurvenschar eine gestrichelte Kurve eingezeichnet. Diese kann näherungsweise mit einer Kurve der Schar verglichen werden, der gemäß Gl. (1) eine Temperatur von 202°C zugeordnet werden müßte. Diese gestrichelte Kurve beschreibt den Verlauf des Richtverhältnisses einer Kennlinie des Oberflächenstromes, die mit plausiblen physikalischen Größen nach der im Anhang angegebenen Formel (6) der Kennlinie des Oberflächenstromes berechnet wurde.

Wesentlich hierbei ist, daß dieser gestrichelten Kurve nicht eine scheinbare Temperatur von 202°C, sondern eine den wirklichen Verhältnissen entsprechende von 27°C zuzuschreiben ist. Dieses Beispiel zeigt wieder, daß experimentelle Strom-Spannungs-Kennlinien sich grundsätzlich durch einen Beitrag des Oberflächenstromes zum Gesamtstrom der Spitzendiode beschreiben lassen.

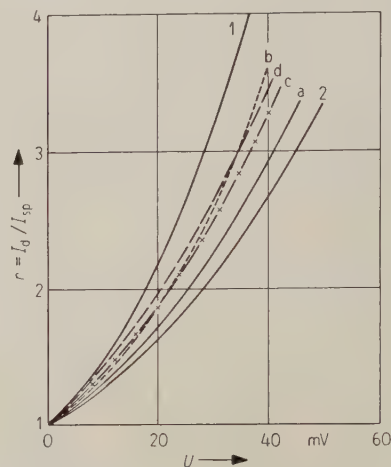


Bild 14. Richtverhältnis in Abhängigkeit von der Spannung. Kurven 1 und 2 ohne, Kurven a, b, c und d mit Berücksichtigung des Oberflächenstromes.

Kurve 1: $\alpha = 38,6$, $t = 27^\circ\text{C}$,
Kurve 2: $\alpha = 24,4$, $t = 202^\circ\text{C}$.

Die Kurven a bis d sind r_{ges} -Kurven der Kennlinien a bis d von Bild 9.

In Bild 14 wird nunmehr gezeigt, daß der Verlauf empirischer Kennlinien in Wirklichkeit nicht auf eine Übertemperatur, sondern auf eine Gesamtheit möglicher Mischungen von Volumen- und Oberflächenströmen zurückgeführt werden kann. Die Kurven 1 und 2 stellen den Verlauf von Richtverhältnissen mit der Spannung gemäß Formel (1) dar. Die Kurve 1 hat einen α -Wert von 38,6, was einer Temperatur von 300°K entspricht, während bei Kurve 2 der Parameter den Wert 24,4 besitzt. (Dies würde eine scheinbare Diodentemperatur von 474°K bedeuten.) Zwischen diesen beiden Kurven 1 und 2 sind Richtverhältniskurven a, b, c und d aufgetragen, die den entsprechend bezeichneten Kennlinien des Bildes 11 zugehörig sind. Hierbei handelt es sich um theoretisch berechnete Kennlinien mit einer Diodentemperatur $T = 300^\circ\text{K}$, die gemäß dem oben erläuterten Modell aus Volumen- und Oberflächen-Stromanteilen resultieren. Nach Bild 11 beschreiben diese theoretischen Kennlinien in befriedigender Weise eine Schar von gemessenen Strom-Spannungs-Charakteristiken. Wenn die theoretischen Kennlinien mit bestimmten physikalischen Bestimmungsgrößen, die weitgehend mit der Erfahrung über die Halbleiteroberfläche in Einklang sind, eine gute Anpassung an den gemessenen Kennlinienverlauf ergeben, so darf man annehmen, daß diese Bestimmungsgrößen den wirklichen Verhältnissen in ausreichender Näherung entsprechen werden. In diesem Sinne repräsentieren die theo-

retischen Kurven a, b, c und d sowohl in Bild 11 als auch in Bild 14 einen realen physikalischen Sachverhalt, der mit dem Experiment weitgehend übereinstimmt.

Es sei an dieser Stelle bemerkt, daß der über den halbkugelförmigen p-n-Übergang fließende Volumenstrom bei Berücksichtigung der Rekombination und Paarbildung im Inneren der raumladungsbehafteten Übergangszone durch einen Rekombinationsstrom vergrößert wird. Bei dem von uns zugrunde gelegten Modell gemäß Bild 8 wurde dieser möglicherweise vorhandene Anteil des Volumenstromes in Gestalt eines zusätzlichen Rekombinationsstromes absichtlich nicht in Betracht gezogen. Hierfür waren mehrere Gründe ausschlaggebend. Erstens ist die durch einen Rekombinationsstromanteil bedingte Vergrößerung des Volumenstromes relativ klein und erreicht bei Germanium den Faktor 2 im allgemeinen nicht, während der Oberflächenstromanteil jedoch ein Mehr- oder Vielfaches des Volumenstromes sein kann. Zweitens sind praktisch die Parameter des Rekombinationsmechanismus in der Übergangszone des halbkugelförmigen p-n-Überganges bei wirklichen Dioden nicht bekannt, so daß man die Zahl freier, unbekannter Parameter des Modells hierdurch vergrößern würde. Der entscheidende Grund für die Vernachlässigung modifizierender Nebeneffekte ist jedoch die Tatsache, daß insbesondere bei Spitzenkontakt-Gleichrichtern der Oberflächenstrom eine experimentell bewiesene Realität darstellt und die wirklichen physikalischen Verhältnisse in den wesentlichen Grundzügen richtig beschreibt.

11. Einfluß der Formierung auf den Diodenstrom

Bei der Herstellung von Spitzenkontakt-Gleichrichtern aus n-leitendem Germanium spielt während des Formierens die Stärke und Zeitdauer des Stromimpulses eine wichtige Rolle. Der Einfluß des Formierprozesses auf die Struktur des Halbleiters in der Nähe der Metallspitze hat verschiedene physikalische Folgen. Von allen diesen Effekten sind jedoch zwei hinsichtlich ihrer Wirkung auf die Diodenkennlinie von primärer Bedeutung, nämlich die Erzeugung einer relativ hohen Leitfähigkeit vom p-Typ innerhalb der halbkugelförmigen Zone und die geometrische Größe dieses formierten Bereiches. Eine Änderung des Radius r_1 durch verschieden starke Formierung beeinflusst sowohl den Volumenstrom als auch den Oberflächenstrom der Diode entsprechend der Änderung der jeweiligen Austrittsflächen. Wesentlicher noch als die Änderung der Größe des Diodenstromes ist jedoch die Änderung der Gestalt der Strom-Spannungs-Kennlinie durch den Einfluß der Formierung. Mit einer geometrischen Veränderung eines Bauelementes sind stets auch Modifikationen der physikalischen Verhältnisse verbunden. Dies kommt bei dem vorliegenden Problem der Spitzendiode in charakteristischer Weise dadurch zum Ausdruck, daß nicht die Größe des Radius r_1 , sondern das Verhältnis von r_1 zur effektiven Diffusionslänge die Kennliniengestalt wesentlich

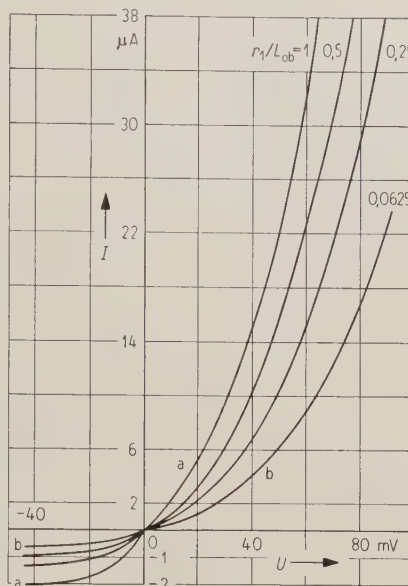


Bild 15. Der berechnete Oberflächenstrom einer Spitzendiode in Abhängigkeit von der angelegten Spannung bei verschiedenen Radien des formierten p-Gebietes; n-Germanium, $E_R = 10^4$ V/cm, $n_D = 10^{15}$ cm $^{-3}$, $\rho = 1,5$ Ω cm, $T = 300^\circ$ K.

lich verändern kann. Hierbei spielt wiederum der Oberflächenstrom in der Anreicherungsschicht eine bemerkenswerte Rolle.

Um diesen Sachverhalt zu demonstrieren, ist in Bild 15 eine berechnete Charakteristik des Oberflächenstromes der Anreicherungsschicht bei 300° K dargestellt. Die Dotierung des n-leitenden Germaniumkristalls beträgt $n_D = 10^{15}$ Atome/cm 3 . Für die elektrische Randfeldstärke wurde ein konventioneller Wert von 10^4 V/cm gewählt (Bild 2). Kurvenparameter ist das Verhältnis des Radius r_1 der formierten halbkugelförmigen p-Zone zur effektiven Diffusionslänge L_{ob} in der Anreicherungsschicht. Die Größe L_{ob} ist ein experimentell zu bestimmender, für den Leitungsmechanismus des Oberflächenstromes charakteristischer Parameter. Er resultiert aus den konventionellen Trägerlebensdauern, Diffusionskonstanten und Beweglichkeiten und beschreibt den ambipolaren Mechanismus der Trägerbewegungen in der Anreicherungsschicht beim Anlegen einer äußeren Spannung (Formel (15)). Der Halbkugelradius r_1 ist vom Formierprozeß der Spitzendiode abhängig, d. h. von Stärke und Form des Stromimpulses, vom Federdruck und vom Halbleitermaterial. Im Zahlenbeispiel des Bildes 15 wurde $r_1 = 25$ μ m und für den Außenradius der Modelldiode $r_2 = 800$ μ m angenommen, was üblichen Richtleiter-Abmessungen entspricht. Die Strom-Spannungscharakteristiken in Bild 15 lassen erkennen, daß der Oberflächenstrom bei festgehaltener Spannung mit größer werdendem Verhältnis r_1/L_{ob} zunimmt. In dieser Gesetzmäßigkeit kommt ersichtlich die Wirkung der Formierung zum Ausdruck.

In diesem Zusammenhang sei noch bemerkt, daß die Gestalt der Kennlinie des Oberflächenstromes

sowie des gesamten Diodenstromes auch durch die anderen Parameter erheblich beeinflusst wird. Bei bestimmten Konfigurationen der Parameter können im Prinzip auch negative Strom-Spannungs-Charakteristiken in bestimmten Spannungsbereichen entstehen.

12. Die effektive Kontaktfläche von Spitzengleichrichtern

Der Verlauf der Potential- bzw. Dichteverteilung in der Modellanreicherungsschicht parallel zur Oberfläche erstreckt sich gemäß der mathematischen Lösung des Randwertproblems über das Oberflächengebiet zwischen den Radien r_1 und r_2 (Bild 8). Die Einwirkung der äußeren Belastung wird jedoch praktisch schon für Entfernungen $r = r^* \ll r_2$ vernachlässigbar klein. Hierdurch wird eine effektive Kontaktfläche der Spitzendiode begrenzt, deren Radius r^* mit der angelegten Spannung zwischen r_1 und r_2 veränderlich ist. Diese durch die Existenz des Oberflächenstromes bedingte spannungsabhängige effektive Kontaktfläche entspricht den wirklichen Verhältnissen bei Spitzendioden weitgehend. Die Vorstellung einer mit der Spannung variierenden Kontaktfläche wurde bereits vor mehreren Jahren von W. SCHOTTKY ausgesprochen¹.

Die Abhängigkeit der effektiven Kontaktfläche von der angelegten Spannung kann im Rahmen des hier zugrunde gelegten Modells wie folgt ermittelt werden. Die Variation der mittleren Defektelektronenkonzentration $\delta\bar{p}$ der Modellanreicherungsschicht als Funktion von r ergibt im Bereich zwischen r_1 und r_2 eine Kurvenschar, für welche die angelegte Spannung Scharparameter ist. Auf Grund der Randbedingungen hat $\delta\bar{p}$ bei $r = r_1$ einen spannungsabhängigen Wert $\delta\bar{p}_{(1)}$ und nimmt von dort mit wachsendem r monoton ab bis zum Randwert $\delta\bar{p}_{(2)} = 0$ bei $r = r_2$. Wir bringen die Kurvenschar $\delta\bar{p}(r)$ mit einer Konstante $\delta\bar{p}_c = \text{const}$ zum Schnitt. Der Wert von $\delta\bar{p}_c$ soll hierbei seinem Betrag nach wesentlich kleiner sein als die Werte $\delta\bar{p}_{(1)}$ der zu einem in Betracht gezogenen Spannungsbereich gehörigen Kurven $\delta\bar{p}(r)$ der Schar. Die auf diese Weise entstehenden Schnittpunkte der Kurvenschar mit der Konstante $\delta\bar{p}_c$ variieren mit der Spannung. Als der zu einer bestimmten Spannung gehörige Radius der effektiven Kontaktfläche soll jetzt derjenige Wert $r = r^*$ definiert werden, für welchen diejenige Kurve der Schar, die dieser Spannung entspricht, mit der Konstante $\delta\bar{p}_c = \text{const}$ zum Schnitt kommt.

Diese Definition erscheint zunächst etwas willkürlich, da die Spannungsabhängigkeit von r^* mit dem Wert $\delta\bar{p}_c$ variiert. Hierbei ist jedoch zu berücksichtigen, daß die so definierte Abschneidevorschrift nach der physikalischen Beschaffenheit der Halbleiteroberfläche gehandhabt werden muß. Der Wert von $\delta\bar{p}_c = \text{const}$ ist die obere Grenze für alle diejenigen Abweichungen der mittleren Konzentra-

tion der Modellanreicherungsschicht, die ihrem Betrag nach so klein sind, daß sie für den Oberflächenstrom praktisch keine Rolle spielen und demzufolge der Einfluß der angelegten Spannung im Bereich $r^* < r \leq r_2$ vernachlässigt werden kann. Insofern stellt r^* eine charakteristische Größe der Halbleiteroberfläche dar.

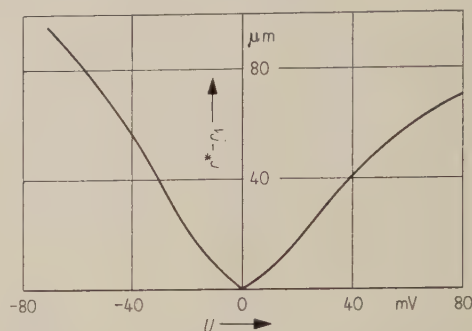


Bild 16. Differenz des Radius r^* der effektiven Kontaktfläche und des Radius r_1 des formierten p-Gebietes in Abhängigkeit von der angelegten Spannung U .

In Bild 16 ist die soeben definierte effektive Kontaktfläche in Abhängigkeit von der angelegten Spannung für ein bestimmtes Beispiel berechnet worden [25]. Nach Voraussetzung sind hierbei die angelegten Spannungen noch so niedrig, daß die elektrische Oberflächenladung bzw. die Randfeldstärke der Inversionsschicht durch die Spannung praktisch nicht modifiziert wird. Als Ordinate ist die Differenz $r^* - r_1$ aufgetragen, die bei einem durch den Formierprozeß bestimmten Wert von r_1 eine spannungsabhängige Ringfläche an der Halbleiteroberfläche charakterisiert. Hierdurch ist der mit der Spannung variierende wirksame Bereich des Oberflächenstromes definitionsgemäß abgegrenzt. Als Beispiel wurden die folgenden konventionellen Werte der Halbleiterparameter verwendet: n-Germanium, Donatorenkonzentration $n_D = 10^{15}$ Atome/cm³, Randfeldstärke der Inversionsschicht $E_R = 10^4$ V/cm, $T = 300^\circ\text{K}$, effektive Diffusionslänge $L_{ob} = 100 \mu\text{m}$, Radius der formierten Halbkugelzone $r_1 = 12,5 \mu\text{m}$, Radius der äußeren Elektrode $r_2 = 800 \mu\text{m}$, Schwellenwert der Variation der mittleren Trägerdichte $\delta\bar{p}_c = 10^{14} \text{cm}^{-3}$. Bild 16 zeigt je eine Kurve für die Sperrichtung und die Flußrichtung. Der Radius der effektiven Kontaktfläche nimmt gemäß der Definition der Variation der mittleren Trägerdichte der Modellanreicherungsschicht in den beiden Belastungsfällen sinngemäß mit wachsendem Betrag der Spannung zu. Bei dieser Zunahme der effektiven Kontaktfläche finden jedoch jeweils bei Fluß- und Sperrbelastung zwei verschiedene Vorgänge in der Inversionsschicht statt. Bei Flußbelastung nimmt die mittlere Trägerkonzentration der Anreicherungsschicht zu, wobei gleichzeitig auch die Dicke der Anreicherungsschicht wächst. Demzufolge findet innerhalb des spannungsabhängigen Ringgebietes zwischen den Radien r^* und r_1 beim Anlegen einer Flußbelastung eine Zunahme der mittleren Trägerkonzentration und so-

¹ Briefliche Diskussionsbemerkungen von W. SCHOTTKY im Jahre 1952 zu einer nicht veröffentlichten Arbeit von E. GROSCHWITZ über den Einfluß von UV-Bestrahlung der Halbleiteroberfläche auf den Gleichrichtereffekt bei Spitzendioden.

mit eine Erhöhung der effektiv wirksamen Leitfähigkeit des Oberflächenstromes statt. Im Sperrfall ist die Tendenz der Veränderungen gerade entgegengesetzt. Die Dicke der Anreicherungsschicht wird mit wachsender Sperrspannung kleiner und ihre mittlere Trägerkonzentration nimmt ab. Diese Verringerung der effektiven, spannungsabhängigen Oberflächenleitfähigkeit erstreckt sich mit zunehmender Sperrspannung über einen ständig größer werdenden Bereich der Oberfläche, d. h., daß durch die effektive Kontaktfläche definierte Einflußgebiet in der Umgebung der Metallspitze dehnt sich aus. Nach Bild 16 wächst die effektive Kontaktfläche mit zunehmender Spannung im Sperrfall etwas stärker als bei Flußbelastung. Auch dieser Sachverhalt ist eine Folge davon, daß die innere Struktur der Anreicherungsschicht bei Fluß- und Sperrbelastung in unsymmetrischer Weise verändert wird.

Ebenso wie bei einer Spitzendiode lassen sich die bisherigen Betrachtungen für den Oberflächenstrom eines legierten Transistors durchführen. Am gravierendsten ist der sogenannte Channelstrom am Collector. Die geometrischen Abmessungen der Inversionsschicht, die den Oberflächenstrom zwischen Collector und Basis führt (Channel), unterscheiden sich jedoch größenordnungsmäßig von der der Spitzendiode, wodurch auch die Größe des Oberflächenstromes beträchtlich verändert wird. Die analytische Beschreibung dieses Sachverhaltes ist gegenüber der Spitzendiode wesentlich komplizierter, weshalb hier der Kürze wegen auf eine Diskussion verzichtet werden muß [19].

Herrn Prof. Dr. G. JUNG danken wir für Diskussionen und Herrn Dipl.-Phys. H. EGER für seine Mitarbeit bei den experimentellen Vorbereitungen und bei der Durchführung zahlreicher Messungen.

Anhang

Die in der vorliegenden Untersuchung verwendeten theoretischen Ergebnisse sind in den Arbeiten [15] bis [19] und [25] ausführlich mitgeteilt und diskutiert worden. Zum Zwecke einer besseren Übersicht soll im folgenden eine kurze Zusammenstellung der in der vorliegenden Arbeit benötigten Formeln gegeben werden.

$$\text{Volumenstrom [18]} \quad I_{\text{vol}} = 2 \pi r_1^2 i_{\text{s vol}} (e^{qU/kT} - 1) \quad (2)$$

$$\text{mit der Sättigungsstromdichte} \quad i_{\text{s vol}} = q \left(\frac{D_n n_p}{L_n \text{vol}} K_n + \frac{D_p p_n}{L_p \text{vol}} K_p \right) \quad (3)$$

und mit den Formfaktoren [18]

$$K_n = \frac{1}{r_1/L_n \text{vol}} \frac{\left(\frac{r_1}{L_n \text{vol}} - 1 \right) \exp \left[\frac{r_1}{L_n \text{vol}} \left(1 - \frac{2r_0}{r_1} \right) \right] + \left(\frac{r_1}{L_n \text{vol}} + 1 \right) \exp \left(-\frac{r_1}{L_n \text{vol}} \right)}{\exp \left[\frac{r_1}{L_n \text{vol}} \left(1 - \frac{2r_0}{r_1} \right) \right] - \exp \left(-\frac{r_1}{L_n \text{vol}} \right)}, \quad (4)$$

$$K_p = \frac{1}{r_1/L_p \text{vol}} \frac{\left(\frac{r_1}{L_p \text{vol}} - 1 \right) \exp \left[-\frac{r_1}{L_p \text{vol}} \left(\frac{2r_2}{r_1} - 1 \right) \right] + \left(\frac{r_1}{L_p \text{vol}} + 1 \right) \exp \left(-\frac{r_1}{L_p \text{vol}} \right)}{\exp \left[-\frac{r_1}{L_p \text{vol}} \left(\frac{2r_2}{r_1} - 1 \right) \right] - \exp \left(-\frac{r_1}{L_p \text{vol}} \right)}. \quad (5)$$

Oberflächenstrom ([15] und [19])

(Der Index (1) bedeutet Abhängigkeit vom äußeren Potential $\varphi_{(1)} = qU/kT$ an der Stelle $r = r_1$.)*

$$I_{\text{ob}} = 2 \pi q x_{(1)}^* \left(\frac{r \partial \delta \bar{p}}{\partial r} \right)_{(1)} \left\{ [\mu_p \bar{p}_0 + (\mu_n + \mu_p) \delta \bar{p}_{(1)}] \left(-\frac{\partial U}{\partial \delta \bar{p}} \right)_{(1)} + (D_n - D_p) \right\}. \quad (6)$$

In Gl. (6) ist die Dicke der Anreicherungsschicht [17]

$$x_{(1)}^* = \frac{1}{\sqrt{K_{1(1)}^*}} \ln \left\{ \frac{1}{Y_{(1)}} \left[(1 + K_{1(1)}^* L_{Dn}^2) + L_{Dn} \sqrt{K_{1(1)}^{*2} L_{Dn}^2 + 2 K_{1(1)}^*} \right] \right\} \quad (7)$$

mit den Strukturparametern [17]

$$K_{1(1)}^* = \frac{2}{L_{Dn}^2} \left(2 \ln \frac{n_D}{n_{(1)}^*} - 1 \right), \quad K_{2(1)}^* = Y_{(1)} = \frac{\left(\frac{du}{dx} \right)_{R(1)}^* L_{Dn} + \sqrt{2} \sqrt{2 \ln \frac{n_D}{n_{(1)}^*} - 1}}{\left(\frac{du}{dx} \right)_{R(1)}^* L_{Dn} - \sqrt{2} \sqrt{2 \ln \frac{n_D}{n_{(1)}^*} - 1}}. \quad (8), (9)$$

Der Einfluß des äußeren Potentials $\varphi_{(1)} = qU/kT$ geht durch die Transformation [17]

$$n_{(1)}^* = n_1 \exp(-\varphi_{(1)}/2) \quad (10)$$

* In der Arbeit [19] ist in Formel (12) bei der Korrektur ein Schreibfehler unbemerkt geblieben. Im zweiten Glied muß in der Klammer vor $1/\tau_p$ an Stelle des positiven ein negatives Vorzeichen stehen.

in die Formeln (7) bis (9) und (16) und folgende ein. In Gl. (6) ist ferner [15], [19]

$$\left(\frac{r \partial \bar{p}}{\partial r}\right)_{(1)} = \frac{r_1}{L_{ob}} \left[\Phi_1 \frac{-i J_1\left(i \frac{r_1}{L_{ob}}\right)}{J_0\left(i \frac{r_1}{L_{ob}}\right)} + \Phi_2 \frac{H_1^{(1)}\left(i \frac{r_1}{L_{ob}}\right)}{i H_0^{(1)}\left(i \frac{r_1}{L_{ob}}\right)} \right] \quad (11)$$

mit den Formfaktoren [15], [19]

$$\Phi_1 = \frac{\delta \bar{p}_{(1)}}{J_0\left(i \frac{r_2}{L_{ob}}\right) i H_0^{(1)}\left(i \frac{r_1}{L_{ob}}\right) - J_0\left(i \frac{r_1}{L_{ob}}\right) i H_0^{(1)}\left(i \frac{r_2}{L_{ob}}\right)}, \quad \Phi_2 = \frac{\delta \bar{p}_{(1)}}{J_0\left(i \frac{r_1}{L_{ob}}\right) i H_0^{(1)}\left(i \frac{r_2}{L_{ob}}\right) - J_0\left(i \frac{r_2}{L_{ob}}\right) i H_0^{(1)}\left(i \frac{r_1}{L_{ob}}\right)} \quad (12), (13)$$

und dem spannungsabhängigen Randwert der Variation der mittleren Trägerdichte der Anreicherungsschicht bei $r = r_1$ [17]

$$\delta \bar{p}_{(1)} = n_D L_{Dn}^2 (K_{1(1)}^* B_{(1)}^* - K_1 B). \quad (14)$$

Die Größe L_{ob} bedeutet die effektive Diffusionslänge der Modellanreicherungsschicht [19]

$$L_{ob} = \sqrt{\frac{D_n + D_p - \frac{\mu_n - \mu_p}{\mu_n + \mu_p} (D_n - D_p)}{\frac{1}{\tau_n} + \frac{1}{\tau_p} - \frac{\mu_n - \mu_p}{\mu_n + \mu_p} \left(\frac{1}{\tau_n} - \frac{1}{\tau_p}\right)}} = \sqrt{\frac{2kT}{q} \frac{\tau \mu_n \mu_p}{\mu_n + \mu_p}}, \quad (\tau_n = \tau_p = \tau). \quad (15)$$

Außerdem gilt in Gl. (6) für den positiven reziproken Wert der Ableitung der Variation der mittleren Trägerdichte nach der Spannung [17]

$$\left(\frac{\partial U}{\partial \bar{p}}\right)_{(1)} = \frac{1}{n_D L_{Dn}^2} (K_{1(1)}^* B_{(1)}^* + K_{1(1)}^* B_{(1)}^*)^{-1}. \quad (16)$$

In Gl. (16) bedeuten die gestrichelten Größen in der rechten Klammer Ableitungen nach der Spannung gemäß Formel (10). Es ist [17]

$$B_{(1)}^* = \frac{\coth\left(\frac{1}{2} \ln Y_{(1)}\right) - \coth\left(\frac{1}{2} \ln A_{(1)}^* + \frac{1}{2} \ln Y_{(1)}\right)}{\ln A_{(1)}^*}, \quad (17)$$

$$A_{(1)}^* = \frac{1}{Y_{(1)}} \left[(1 + K_{1(1)}^* L_{Dn}^2) + L_{Dn} \sqrt{K_{1(1)}^* L_{Dn}^2 + 2 K_{1(1)}^*} \right]. \quad (18)$$

Schrifttum

- [1] HEILAND, G., Zur Theorie der Anreicherungsrandschicht an der Oberfläche von Halbleitern. Z. Phys. **138** [1954], 459–464, **142** [1955], 415–432, **148** [1957], 15–27.
- [2] JÄNTSCH, O., Persönliche Mitteilung einer Arbeit „Das Sperrverhalten von Siliziumgleichrichtern in feuchten Gasen“; (z. Zt. im Druck in Z. Naturforsch.).
- [3] BARDEEN, J. und BRATTAIN, W. H., Physical principles involved in transistor action. Phys. Rev. **75** [1949], 1208–1225.
- [4] AIGRAIN, P., Contribution à l'étude des phénomènes de redressement et de transistance dans le germanium. Ann. Phys. (Paris) **7** [1952], 140–184.
- [5] BROWN, W. L., n-type surface conductivity on p-type germanium. Phys. Rev. **91** [1953], 518–527.
- [6] CHRISTENSEN, H., Surface "channel" detection method for semiconductors. Phys. Rev. **96** [1954], 827–828, **98** [1955], 1766–1775.
- [7] MCWHORTER, A. L. und KINGSTON, R. H., Channels and excess reverse current in grown germanium p-n junction diodes. Proc. Inst. Radio Engrs. **42** [1954], 1376–1380.
- [8] KINGSTON, R. H., Water-vapor-induced n-type surface conductivity on p-type germanium. Phys. Rev. **98** [1955], 1766–1775.
- [9] DEMARS, G. A., STATZ, H. und DAVIS, L., Measurement and interpretation of conductance of p-type inversion layers on germanium. Phys. Rev. **98** [1955], 539–540.
- [10] CUTLER, M. und BATH, H. M., Measurement of surface recombination velocity in silicon by steady-state photoconductance. Proc. Inst. Radio Engrs. **45** [1957], 39–43.
- [11] GUBANOW, A. I., J. exp. theor. Phys. UdSSR **22** [1952], 204.
- [12] GARRETT, C. G. B. und BRATTAIN, W. H., Physical theory of semiconductor surfaces. Phys. Rev. **99** [1955], 376–388.
- [13] STATZ, H., DEMARS, G. A., DAVIS, L. und ADAMS, A., Surface states on silicon and germanium surfaces. Phys. Rev. **101** [1956], 1272–1281.
- [14] HARTEN, H. U. und SCHULTZ, W., Referat in „W. SCHOTTKY, Halbleiterprobleme III“, S. 76, Verlag F. Vieweg, Braunschweig 1956. HARTEN, H. U., KOCH, W., RATH, H. L. und SCHULTZ, W., Die Sperrschichtkapazität des legierten Germanium-Indium-Gleichrichters. Z. Phys. **138** [1954], 336–344.
- [15] SCHULTZ, W., Zur Theorie der Gleichrichtung am Kontakt Metall-Halbleiter. Z. Phys. **138** [1954], 598–612.
- [16] GROSCHWITZ, E., Elektrische Eigenschaften von Inversions-schichten an Halbleiteroberflächen. Solid State Physics in Electronics and Telecommunications (Published by Academic Press Inc., London 1960, Vol. 1, 575–586).
- [17] GROSCHWITZ, E., HOFMEISTER, E. und EBHARDT, R., Zur Physik des Spitzengleichrichters. Vortrag auf der Physikertagung in Berlin, 29. Sept. bis 3. Okt. 1959.
- [18] GROSCHWITZ, E. und EBHARDT, R., Zur Theorie der Inversions-schichten an Halbleiteroberflächen. Z. angew. Phys. **11** [1959], 9–19.
- [19] HOFMEISTER, E. und GROSCHWITZ, E., Der Einfluß der geometrischen und physikalischen Verhältnisse an der Spitze von Germaniumdioden auf die Kennlinie. Z. angew. Phys. **10** [1958], 109–114.
- [20] GROSCHWITZ, E. und EBHARDT, R., Zur Theorie der Oberflächenströme an Halbleiteroberflächen. Z. angew. Phys. **11** [1959], 296–304.
- [21] STATZ, H., DEMARS, G. A., jr. und ADAMS, A., jr., Beitrag in „Semiconductor Surface Physics“, herausgegeben von R. H. KINGSTON, Philadelphia, University of Pennsylvania Press 1957.
- [22] SCHRIEFFER, J. R., Effective carrier mobility in surface-space charge layers. Phys. Rev. **97** [1955], 641–646.
- [23] THEDECK, R., Beitrag zur Kenntnis des Mechanismus von Punkt-kontakt-Transistoren. Z. angew. Phys. **5** [1953], 163–165, 165–166.
- [24] BENEKING, H., Zum Gleich- und Wechselstromverhalten von Punkt-kontaktdioden. Z. angew. Phys. **9** [1957], 626–631, **10** [1958], 216–225.
- [25] BETHE, H. A., Theory of the boundary layer of crystal rectifiers. RL-Report No. 43–12, 23. 11. 1942.
- [26] GROSCHWITZ, E. und EBHARDT, R., Die effektive Kontaktfläche von Spitzenkristallgleichrichtern. Z. angew. Phys. **11** [1959], 342–346.

Berechnung der Verzerrungs- und Störeffekte bei Transistor-Verstärkerstufen auf Grund des Ersatzschaltbildes

Von JAKOB S. VOGEL und MAX J. O. STRUTT

Mitteilung aus dem Institut für höhere Elektrotechnik der Eidgenössischen Technischen Hochschule Zürich

A.E.Ü. 14 [1960], 397–404; eingegangen am 15. Juli 1960)

DK 621.375.4:621.391.832.4

Mit Hilfe einer zweckmäßigen Approximation der Eingangs- bzw. Übertragungscharakteristik des Transistors durch eine Summe von Exponentialfunktionen können die nichtlinearen Verzerrungseffekte wie harmonische Verzerrungen, Modulationsverzerrung, Modulationssteigerung und Kreuzmodulation für Flächentransistoren einfach bestimmt werden. Die Berechnungen erfolgen für das Niederfrequenzgebiet für verschiedene Lastwiderstände und Zuleitungswiderstände im Basiskreis bei beliebig großen (erlaubten) Eingangssignalen.

Die Berechnungen werden auf zweistufige Verstärkerschaltungen erweitert und der Übergang zur Bestimmung der Verzerrungen bei hohen Frequenzen wird gegeben.

Using an appropriate approximation of the input- and of the transfer-characteristic of the transistor by a sum of exponential functions, non-linear distortion effects such as the generation of harmonics, modulation rise, distortion of modulation and cross-modulation of junction transistors may be evaluated easily. The calculations are carried out in the low frequency range for various load resistances and lead resistances of the base circuit for input signals varying up to the largest permissible values.

The calculations are applied to two-stage amplifier circuits. The extension of the theory, necessary to evaluate distortion at high frequencies is indicated.

1. Einleitung

In letzter Zeit sind verschiedene Arbeiten erschienen [1]–[4], welche gewisse nichtlineare Verzerrungen in Transistorschaltungen besonders für kleine Signale und niedrige Frequenzen behandeln. Die vorliegende Arbeit soll eine einfache Berechnungsart zeigen, die es ermöglicht, harmonische Verzerrungen, Modulationsverzerrungen, Kreuzmodulation und Mischung für beliebige, zulässige Signalamplituden direkt aus dem Transistorersatzschaltbild zu bestimmen.

2. Der Eingangswiderstand R_i bei kurzgeschlossenem Ausgang und niedrigen Frequenzen

In der Emitterschaltung gilt für kleine Signalamplituden mit genügender Genauigkeit das in

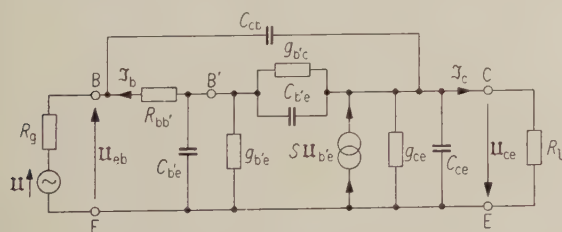


Bild 1. Hochfrequenz-Ersatzschaltung für einen Flächentransistor in Emitterschaltung.

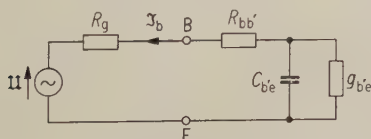


Bild 2. Hochfrequenz-Ersatzschaltung für einen Flächentransistor in Emitterschaltung bei kurzgeschlossenem Ausgang.

Bild 1 dargestellte Transistor-HF-Ersatzschema [5], [6], das sich bei ausgangseitigem Kurzschluß auf dasjenige von Bild 2 vereinfacht. Der Emitterleitwert $g_{b'e}$ ist dabei dem Collectorgleichstrom I_{0C} im betreffenden Arbeitspunkt direkt proportional¹ [5] bis [7]:

$$g_{b'e} = \frac{e}{kT} \frac{I_{0C}}{\alpha_{fe0}} = \frac{I_{0C}}{U_T \alpha_{fe0}} \quad (1)$$

mit der Abkürzung

$$kT/e = U_T \text{ „Temperaturspannung“,} \quad (2)$$

k Boltzmannsche Konstante,

e Betrag der Elementarladung,

T absolute Temperatur.

Für 300°K wird $U_T = 25,9$ mV.

$\alpha_{fe0} = di_C/di_B$ ist die Kurzschluß-Stromverstärkung in Emitterschaltung bei tiefen Frequenzen. Die in Bild 1 eingeführten Bezeichnungen 1 und 3 gelten nur für einwellige Spannungen und Ströme.

Für größere Signalamplituden und Wechselströme in der Größenordnung von I_{0C} sind für die weiteren Berechnungen die Momentanwerte (im folgenden mit dem Index m bezeichnet) der Elemente des Ersatzschaltbildes zu berücksichtigen. Besonders interessiert der Momentanwert des Emitterleitwertes $g_{b'e m}$, der von $g_{b'e}$ (Bild 1) erheblich abweichen kann.

Gleichung (1) geht über in

$$g_{b'e m} = \frac{i_C}{U_T \alpha_{fe0 m}} \quad (1')$$

Der Momentanwert i_C des Collectorstromes setzt sich aus zwei Anteilen zusammen: aus dem Gleichstrom I_{0C} im betreffenden Arbeitspunkt und aus dem Momentanwert i_c , der durch eine Eingangs-

¹ Eine Liste der verwendeten Symbole findet sich am Schluß der Arbeit.

Wechselspannung $u(t)$ erzeugt wird. Wegen der Gleichrichtereigenschaft der Eingangskennlinie ist in i_c ebenfalls ein Gleichstromanteil enthalten.

Es gilt also

$$i_c = I_{0C} + i_c. \quad (3)$$

Für die folgende Rechnung soll vorläufig angenommen werden, α_{fe0} sei stromunabhängig (siehe Abschnitt 4). Dies gilt auch für den Basiszuleitungswiderstand $R_{bb'}$. Somit ist $\alpha_{fe0m} = \alpha_{fe0}$ und $R_{bb'm} = R_{bb'}$. Solange man weit unterhalb der Grenzfrequenz arbeitet, darf die Wirkung der Emittierkapazität $C_{b'e}$ bzw. $C_{b'e m}$ vernachlässigt werden. Faßt man den Generatorinnenwiderstand R_g mit $R_{bb'}$ zu R_b zusammen, so wird der momentane Eingangswiderstand bei kurzgeschlossenem

$$\text{Ausgang } R_{1m} = R_b + \frac{1}{g_{b'e m}}.$$

Bezeichnet man die Momentanwerte von Eingangsspannung und Basisstrom mit u bzw. i_B , so folgt mit Gl. (1')

$$R_{1m} = \frac{du}{di_B} = \frac{du}{di_C} \frac{di_C}{di_B} = R_b + \frac{\alpha_{fe0} U_T}{i_C} \quad (4)$$

und mit $\frac{di_C}{di_B} = \alpha_{fe0m} = \alpha_{fe0},$

$$\frac{du}{di_C} = \frac{R_b}{\alpha_{fe0}} + \frac{U_T}{i_C}. \quad (5)$$

Es ist zu beachten, daß die Pfeilrichtungen von i_C , i_c und I_{0C} bzw. i_B mit den in Bild 1 angegebenen Richtungen für \mathfrak{I}_c bzw. \mathfrak{I}_b übereinstimmen, ebenso der Pfeilsinn von u bzw. $u(t)$ mit demjenigen von U .

Die Lösung der Differentialgleichung mit den Anfangsbedingungen $i_C = I_{0C}$, $u = 0$ ergibt

$$u = \frac{R_b i_c}{\alpha_{fe0}} + U_T \ln \frac{i_c}{I_{0C}},$$

$$\frac{i_c}{I_{0C}} = \exp \left(\frac{u - \frac{R_b i_c}{\alpha_{fe0}}}{U_T} \right). \quad (6)$$

Wegen der Nichtlinearität der Funktion $i_c = f(u)$ entstehen die Verzerrungen in Transistorschaltungen.

3. Die nichtlinearen Verzerrungen

Für die folgenden Berechnungen machen wir von den Entwicklungsformeln

$$\exp(x \sin \omega t) = \quad (7)$$

$$= I_0(x) + 2 \sum_{m=1}^{\infty} (-1)^m I_{2m}(x) \cos 2m\omega t +$$

$$+ 2 \sum_{m=0}^{\infty} (-1)^m I_{2m+1}(x) \sin (2m+1)\omega t$$

und

$$\exp(x \cos \omega t) = I_0(x) + 2 \sum_{m=1}^{\infty} I_m(x) \cos m\omega t \quad (8)$$

Gebrauch, wobei $I_m(x)$ die modifizierte Besselsche Funktion m -ter Ordnung mit dem Argument x bezeichnet. In der Umgebung des Nullpunktes läßt sich diese Funktion in eine Taylorreihe entwickeln [8]:

$$I_m(x) = \frac{1}{m!} \left(\frac{x}{2} \right)^m + \frac{1}{1!(m+1)!} \left(\frac{x}{2} \right)^{m+2} +$$

$$+ \frac{1}{2!(m+2)!} \left(\frac{x}{2} \right)^{m+4} + \quad (9)$$

$$+ \frac{1}{3!(m+3)!} \left(\frac{x}{2} \right)^{m+6} + \dots$$

Für die Ordnungen $m = 0, 1, 2$ und 3 ergibt sich somit

$$I_0(x) = 1 + \frac{x^2}{4} + \frac{x^4}{64} + \frac{x^6}{2304} + \dots, \quad (10)$$

$$I_1(x) = \frac{x}{2} + \frac{x^3}{16} + \frac{x^5}{384} + \frac{x^7}{18432} + \dots, \quad (11)$$

$$I_2(x) = \frac{x^2}{8} + \frac{x^4}{96} + \frac{x^6}{3072} + \dots, \quad (12)$$

$$I_3(x) = \frac{x^3}{48} + \frac{x^5}{768} + \frac{x^7}{30720} + \dots. \quad (13)$$

Im weiteren benötigen wir noch die Symmetrieformel

$$I_m(-x) = (-1)^m I_m(x). \quad (14)$$

Wird am Eingang des Transistors eine sinusförmige Spannung $u(t) = u_1 = \hat{U}_1 \sin \omega t$ angelegt und bezeichnet man den Quotienten \hat{U}_1/U_T mit x , so schreibt sich Gl. (6)

$$\frac{i_c}{I_{0C}} = \exp \left(\frac{-R_b i_c}{\alpha_{fe0} U_T} \right) \exp(x \sin \omega t). \quad (15)$$

Wegen des Faktors $\exp \left(\frac{-R_b i_c}{\alpha_{fe0} U_T} \right)$ kann die Entwicklungsformel (7) nicht direkt auf Gl. (15) angewendet werden. Wir nähern deshalb Gl. (15) durch die Summe von $n+1$ Exponentialfunktionen der Form

$$\frac{i_c}{I_{0C}} = \sum_{i=1}^{n+1} a_i \exp \left(k_i \frac{u}{U_T} \right) = \sum_{i=1}^{n+1} a_i \exp(k_i x \sin \omega t) \quad (16)$$

an und bestimmen die Koeffizienten a_i so, daß die durch Gl. (15) und (16) dargestellten Kennlinien $i_c = f(u)$ sowohl im Arbeitspunkt I_{0C} , als auch in den ersten n Ableitungen $d^n i_c / du^n$ im Arbeitspunkt übereinstimmen. Diese Bestimmung gilt für beliebige Werte von k_i .

Aus Gl. (6) bestimmen wir nun die Ableitungen $d^n i_c / du^n$ an der Stelle $i_c = I_{0C}$:

$$\left. \frac{di_c}{du} \right|_{i_c=I_{0C}} = \frac{I_{0C}}{U_T} \frac{1}{1+q} = \frac{I_{0C}}{U_T} \alpha, \quad (17)$$

$$\left. \frac{d^2 i_c}{du^2} \right|_{i_c=I_{0C}} = \frac{I_{0C}}{U_T^2} \frac{1}{(1+q)^3} = \frac{I_{0C}}{U_T^2} \beta, \quad (18)$$

$$\left. \frac{d^3 i_c}{du^3} \right|_{i_c=I_{0C}} = \frac{I_{0C}}{U_T^3} \frac{1-2q}{(1+q)^5} = \frac{I_{0C}}{U_T^3} \gamma, \quad (19)$$

und analog ergeben die weiteren Ableitungen

$$\frac{I_{0C}}{U_T^4} \delta, \quad \frac{I_{0C}}{U_T^5} \epsilon, \quad \frac{I_{0C}}{U_T^6} \zeta, \quad \frac{I_{0C}}{U_T^7} \eta, \quad \frac{I_{0C}}{U_T^8} \vartheta \text{ usw.},$$

wobei die „reduzierten Ableitungen“ $\delta, \varepsilon, \zeta, \eta, \vartheta$ die Werte

$$\delta = \frac{1 - 8q + 6q^2}{(1 + q)^7}, \quad (20)$$

$$\varepsilon = \frac{1 - 22q + 58q^2 - 24q^3}{(1 + q)^9}, \quad (21)$$

$$\zeta = \frac{1 - 52q + 328q^2 - 444q^3 + 120q^4}{(1 + q)^{11}}, \quad (22)$$

$$\eta = \frac{1 - 114q + 1452q^2 - 4400q^3 + 3708q^4 - 720q^5}{(1 + q)^{13}}, \quad (23)$$

$$\vartheta = \frac{1 - 240q + 5610q^2 - 32120q^3 + 58140q^4 - 33984q^5 + 5040q^6}{(1 + q)^{15}} \quad (24)$$

besitzen. Es wurde dabei in allen Ableitungen die Abkürzung

$$q = \frac{I_{0C} R_b}{U_T \alpha_{fe0}} \quad (25)$$

eingeführt. Die Verläufe der reduzierten Ableitungen sind aus den Bildern 3, 4 und 5 ersichtlich.

Mit den früher angeführten Bedingungen für die Koeffizienten a_i erhält man aus Gl. (16) das Gleichungssystem

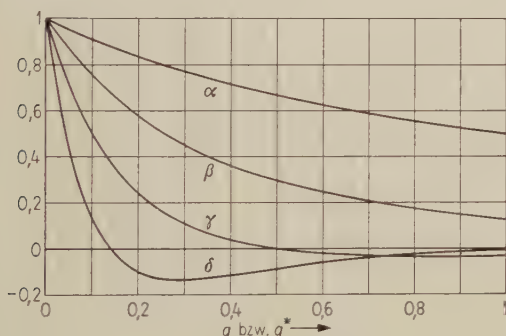


Bild 3. Verlauf der reduzierten Ableitungen $\alpha, \beta, \gamma, \delta$ als Funktion von $q = R_b I_{0C} / U_T \alpha_{fe0}$ bzw. $q^* = R_b^* I_{0C} / U_T^* \alpha_{fe0}$.

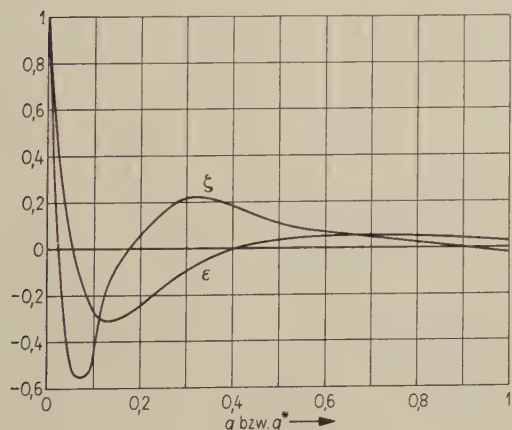


Bild 4. Verlauf der reduzierten Ableitungen ε, ζ als Funktion von $q = R_b I_{0C} / U_T \alpha_{fe0}$ bzw. $q^* = R_b^* I_{0C} / U_T^* \alpha_{fe0}$.

$$\begin{aligned} \sum_{i=1}^{n+1} a_i &= 1 \\ \sum_{i=1}^{n+1} a_i k_i &= \alpha \\ \sum_{i=1}^{n+1} a_i k_i^2 &= \beta \\ \sum_{i=1}^{n+1} a_i k_i^3 &= \gamma \\ &\dots \\ \sum_{i=1}^{n+1} a_i k_i^n &= \nu, \end{aligned} \quad (26)$$

worin ν die n -te reduzierte Ableitung sein soll. Hieraus könnte man die Koeffizienten a_i für beliebige k_i bestimmen. Wie gleich gezeigt werden soll, ist dies jedoch nicht notwendig.

3.1. Die harmonischen Verzerrungen

3.1.1. Die Grundharmonische

Auf die Form (16) läßt sich nun die Entwicklung nach Gl. (7) anwenden. Die Grundharmonische des Collectorstromes ergibt sich für $m = 0$

$$\frac{i_{c1}}{I_{0C}} = 2 \sum_{i=1}^{n+1} a_i I_1(k_i x) \sin \omega t.$$

Mit Gl. (11) erhält man

$$\begin{aligned} \frac{i_{c1}}{I_{0C}} = & \left(2 \sum_{i=1}^{n+1} a_i \frac{k_i x}{2} + 2 \sum_{i=1}^{n+1} a_i \frac{k_i^3 x^3}{16} + \right. \\ & \left. + 2 \sum_{i=1}^{n+1} a_i \frac{k_i^5 x^5}{384} + 2 \sum_{i=1}^{n+1} a_i \frac{k_i^7 x^7}{18432} + \dots \right) \sin \omega t, \end{aligned}$$

und mit Gl. (26) lassen sich die Summen umschreiben:

$$\frac{i_{c1}}{I_{0C}} = \left(\alpha x + \gamma \frac{x^3}{8} + \varepsilon \frac{x^5}{192} + \eta \frac{x^7}{9216} + \dots \right) \sin \omega t. \quad (27)$$

Wo die Entwicklung abgebrochen werden darf, hängt von der Größe von x und von q ab. Beispiels-

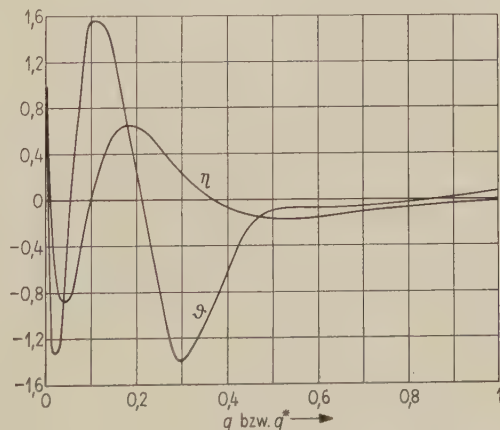


Bild 5. Verlauf der reduzierten Ableitungen η, ϑ als Funktion von $q = R_b I_{0C} / U_T \alpha_{fe0}$ bzw. $q^* = R_b^* I_{0C} / U_T^* \alpha_{fe0}$.

weise erhält man für $x = 1$ und $q = 0,1$ (z. B. $R_b = 100 \Omega$, $I_{0C} = 2,59 \text{ mA}$, $\alpha_{te0} = 100$, $U_T = 25,9 \text{ mV}$)

$$i_{c1}/I_{0C} = (0,91 + 0,06 - 0,001) \sin \omega t.$$

Der 3. Term ist schon vernachlässigbar.

Man sieht aus der Herleitung für die Grundharmonische, daß die Koeffizienten a_i und k_i nicht von Interesse sind, da nach dem Einsetzen der Entwicklungsformeln für die Besselschen Funktionen immer nur die reduzierten Ableitungen auftreten.

3.1.2. Die zweite Harmonische

Für $m = 1$ erhält man aus Gl. (7) und (16)

$$\frac{i_{c2}}{I_{0C}} = -2 \sum_{i=1}^{n+1} a_i I_2(k_i x) \cos 2 \omega t$$

und mit Gl. (12) und (26) (28)

$$\frac{i_{c2}}{I_{0C}} = -\left(\beta \frac{x^2}{4} + \delta \frac{x^4}{48} + \zeta \frac{x^6}{1536} + \dots\right) \cos 2 \omega t.$$

3.1.3. Die dritte Harmonische

Die analoge Berechnung ergibt (29)

$$\frac{i_{c3}}{I_{0C}} = -\left(\gamma \frac{x^3}{24} + \varepsilon \frac{x^5}{384} + \eta \frac{x^7}{15360} + \dots\right) \sin 3 \omega t.$$

Man ersieht aus der Formel sofort, daß für $x \ll 1$ die 3. Harmonische für $\gamma = 0$ verschwindet:

$$\gamma = \frac{1 - 2q}{(1 + q)^5} = 0, \quad q = 0,5. \quad (30)$$

Für größere Werte von x müssen die höheren Ableitungen berücksichtigt werden. Der Nullpunkt verschiebt sich dann etwas, für $x = 1$ wird $q = 0,51$, und für $x = 2$ wird $q = 0,53$.

Die harmonischen Verzerrungen sind nun durch die Amplitudenverhältnisse von Oberwellen zur Grundwelle bestimmt:

$$\frac{A_2}{A_1} = \frac{\beta \frac{x}{4} + \delta \frac{x^3}{48} + \zeta \frac{x^5}{1536} + \dots}{\alpha + \gamma \frac{x^2}{8} + \varepsilon \frac{x^4}{192} + \dots}, \quad (31)$$

$$\text{für } x \ll 1 \text{ gilt} \quad \frac{A_2}{A_1} = \frac{\beta}{\alpha} \frac{x}{4}; \quad (31')$$

$$\frac{A_3}{A_1} = \frac{\gamma \frac{x^2}{24} + \varepsilon \frac{x^4}{384} + \dots}{\alpha + \gamma \frac{x^2}{8} + \varepsilon \frac{x^4}{192} + \dots}, \quad (32)$$

$$\text{für } x \ll 1 \text{ gilt} \quad \frac{A_3}{A_1} = \frac{\gamma}{\alpha} \frac{x^2}{24}. \quad (32')$$

3.2. Die Verzerrungen der Modulation

Wird am Transistoreingang ein einfaches, mit der Kreisfrequenz p amplitudenmoduliertes Signal der Form $u_1 = \hat{U}_1(1 + M \cos pt) \sin \omega t$ angelegt, so erhält man mit den Abkürzungen $\hat{U}_1/U_T = x$ und $\hat{U}_1(1 + M \cos pt)/U_T = x_p$ für den Strom mit der Kreisfrequenz ω den zu Gl. (27) analogen Wert

$$\frac{i_{c1}}{I_{0C}} = \left(\alpha x_p + \gamma \frac{x_p^3}{8} + \dots\right) \sin \omega t,$$

was mit

$$\begin{aligned} x_p^3 &= x^3 (1 + M \cos pt)^3 = \\ &= x^3 \left[\left(1 + \frac{3}{2} M^2\right) + \left(3M + \frac{3}{4} M^3\right) \cos pt + \right. \\ &\quad \left. + \frac{3}{2} M^2 \cos 2pt + \frac{M^3}{4} \cos 3pt \right] \end{aligned} \quad (33)$$

auf

$$\begin{aligned} \frac{i_{c1}}{I_{0C}} &= x \left\{ \left[\alpha + \frac{x^2 \gamma}{8} \left(1 + \frac{3}{2} M^2\right) \right] + \right. \\ &\quad \left. + M \left[\alpha + \frac{3x^2 \gamma}{8} \left(1 + \frac{M^2}{4}\right) \right] \cos pt + \right. \\ &\quad \left. + \frac{3x^2 \gamma M^2}{16} \cos 2pt + \frac{x^2 \gamma M^3}{32} \cos 3pt \right\} \sin \omega t \end{aligned} \quad (34)$$

führt.

Der neue Modulationsgrad ergibt sich so zu

$$M^* = \frac{\alpha + \frac{3x^2 \gamma}{8} \left(1 + \frac{M^2}{4}\right)}{\alpha + \frac{x^2 \gamma}{8} \left(1 + \frac{3}{2} M^2\right)} M, \quad (35)$$

und die relative Änderung des Modulationsgrades wird

$$\frac{M^* - M}{M} = \frac{\frac{x^2 \gamma}{4} \left(1 - \frac{3}{8} M^2\right)}{\alpha + \frac{x^2 \gamma}{8} \left(1 + \frac{3}{2} M^2\right)}, \quad (36)$$

und für $x \ll 1$ und $M < 1$ gilt

$$\frac{M^* - M}{M} = \frac{\gamma}{\alpha} \frac{x^2}{4}. \quad (36')$$

Die relativen Verzerrungen der Modulation, ausgedrückt durch die Definitionsgleichung

$$d_n = \frac{\text{Amplitude der Modulation mit der Kreisfrequenz } np}{\text{Amplitude der Modulation mit der Kreisfrequenz } p}$$

werden:

$$d_2 = \frac{3x^2 \gamma M}{16\alpha + 6x^2 \gamma \left(1 + \frac{M^2}{4}\right)}, \quad (37)$$

für $x \ll 1$, $M < 1$ gilt

$$d_2 = \frac{\gamma}{\alpha} \frac{3}{16} x^2 M; \quad (37')$$

$$d_3 = \frac{x^2 \gamma M^2}{32\alpha + 12x^2 \gamma \left(1 + \frac{M^2}{4}\right)}, \quad (38)$$

für $x \ll 1$, $M < 1$ gilt

$$d_3 = \frac{\gamma}{\alpha} \frac{x^2 M^2}{32}. \quad (38')$$

Für sehr große Werte von x müssen weitere Glieder der Taylorentwicklung (27) berücksichtigt werden, so daß dann die höheren Ableitungen ε , η , ... auftreten.

Man ersieht aus Gl. (36), (37) und (38), daß bei $x \ll 1$ die Modulationsverzerrungen für $\gamma = 0$, also $q = 0,5$, verschwinden.

3.3. Die Kreuzmodulation

Erscheint am Eingang neben dem Nutzsignal $u_1 = \hat{U}_1 \sin \omega_1 t$ noch ein Störsignal $u_2 = \hat{U}_2(1 + M \cos p t) \sin \omega_2 t$, also $u(t) = u_1 + u_2$, so wird das Nutzsignal mit der Kreisfrequenz p moduliert. Verwendet man die Abkürzungen

$$\frac{\hat{U}_1}{U_T} = x, \quad \frac{\hat{U}_2}{U_T} = z, \quad \frac{\hat{U}_2(1 + M \cos p t)}{U_T} = z_p,$$

so schreibt sich Gl. (16) jetzt

$$\begin{aligned} \frac{i_C}{I_{0C}} &= \sum_{i=1}^{n+1} a_i \exp\left(k_i \frac{u}{U_T}\right) = \\ &= \sum_{i=1}^{n+1} a_i \exp(k_i x \sin \omega_1 t) \exp(k_i z_p \sin \omega_2 t). \end{aligned}$$

Wendet man auf die Exponentialfunktionen die Entwicklung (7) an, so ergibt sich

$$\begin{aligned} \frac{i_C}{I_{0C}} &= \\ \sum_{i=1}^{n+1} a_i &\left[I_0(k_i x) + 2 \sum_{m=1}^{\infty} (-1)^m I_{2m}(k_i x) \cos 2m \omega_1 t + \right. \\ &+ 2 \sum_{m=0}^{\infty} (-1)^m I_{2m+1}(k_i x) \sin (2m+1) \omega_1 t \Big] \cdot \\ &\cdot \left[I_0(k_i z_p) + 2 \sum_{m=1}^{\infty} (-1)^m I_{2m}(k_i z_p) \cos 2m \omega_2 t + \right. \\ &+ 2 \sum_{m=0}^{\infty} (-1)^m I_{2m+1}(k_i z_p) \sin (2m+1) \omega_2 t \Big]. \end{aligned} \quad (39)$$

Es interessiert uns jedoch nur der modulierte Strom mit der Kreisfrequenz ω_1 . Gl. (39) vereinfacht sich zu

$$\frac{i_C}{I_{0C}} = \sum_{i=1}^{n+1} a_i \cdot 2 \cdot I_1(k_i x) I_0(k_i z_p) \sin \omega_1 t \quad (40)$$

und durch Anwendung von Gl. (10) und (11) zu

$$\begin{aligned} \frac{i_C}{I_{0C}} &= \left(2 \sum_{i=1}^{n+1} a_i \frac{k_i x}{2} + 2 \sum_{i=1}^{n+1} a_i \frac{k_i^3 x^3}{16} + \dots + \right. \\ &+ 2 \sum_{i=1}^{n+1} a_i \frac{k_i x}{2} \frac{k_i^2 z_p^2}{4} + \\ &+ 2 \sum_{i=1}^{n+1} a_i \frac{k_i^3 x^3}{16} \frac{k_i^2 z_p^2}{4} + \dots \Big) \sin \omega_1 t, \end{aligned}$$

und mit Gl. (26) gilt dann

$$\begin{aligned} \frac{i_C}{I_{0C}} &= \left[\alpha x + \gamma \frac{x^3}{8} + \dots + \right. \\ &+ \left(\gamma x + \varepsilon \frac{x^3}{8} + \eta \frac{x^5}{192} + \dots \right) \frac{z_p^2}{4} \Big] \sin \omega_1 t. \end{aligned} \quad (41)$$

Auch hier ist die Zahl der zu berücksichtigenden Glieder von x abhängig.

Die Kreuzmodulation rührt vom Term mit z_p^2 her. Soll sie verschwinden, so muß

$$\gamma + \varepsilon \frac{x^2}{8} + \eta \frac{x^4}{192} + \dots = 0 \quad (42)$$

gelten, was für $x \ll 1$ wieder auf $\gamma = 0$ führt. Be-

rücksichtigt man die drei ersten Terme, so erhält man für den Nullpunkt der Kreuzmodulation eine Abhängigkeit des Wertes q von der Aussteuerung. Der Verlauf nach Bild 6 dürfte bis $x = 3$ gelten. Für größere Werte müssen noch zusätzliche Ableitungen berücksichtigt werden. Möglicherweise treten dann noch weitere Nullpunkte für andere Werte von q auf, doch wird dies am besten experimentell untersucht.

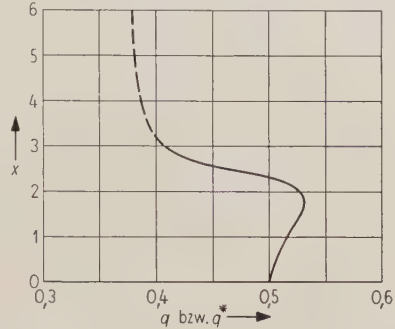


Bild 6. Nullpunkt der Kreuzmodulation für verschiedene Werte von x (gültig für $x \leq 3$).

Führt man die Entwicklung

$$z_p^2 = z^2(1 + M \cos p t)^2 = \quad (43)$$

$$= z^2 \left[\left(1 + \frac{M^2}{2} \right) + 2 M \cos p t + \frac{M^2}{2} \cos 2 p t \right]$$

in Gl. (41) ein, so ergibt sich

$$\begin{aligned} \frac{i_C}{I_{0C}} &= \left\{ \alpha x + \gamma \frac{x^3}{8} + \frac{x z^2}{4} \left(1 + \frac{M^2}{2} \right) \left(\gamma + \varepsilon \frac{x^2}{8} \right) + \right. \\ &+ \frac{x z^2}{2} M \left(\gamma + \varepsilon \frac{x^2}{8} \right) \left(\cos p t + \frac{M}{4} \cos 2 p t \right) \Big\} \sin \omega_1 t. \end{aligned} \quad (44)$$

Der Kreuzmodulationsgrad wird angenähert

$$M_k = \frac{\frac{z^2}{2} M \left(\gamma + \varepsilon \frac{x^2}{8} \right)}{\alpha + \gamma \frac{x^2}{8} + \frac{z^2}{4} \left(1 + \frac{M^2}{2} \right) \left(\gamma + \varepsilon \frac{x^2}{8} \right)}, \quad (45)$$

und für $z \ll 1$, $x \ll 1$ ergibt sich

$$M_k = \frac{\gamma}{\alpha} \frac{z^2 M}{2}. \quad (45')$$

Vergleicht man Gl. (32'), (36'), (37'), (38') und (45'), so sieht man, daß alle diese Verzerrungserscheinungen für $x \ll 1$ eng miteinander verknüpft sind, und daß es im Prinzip genügt, eine davon zu messen und die anderen daraus zu berechnen.

Aus Gl. (45) kann noch die effektive Störspannung für 1% Kreuzmodulation, d. h. $M_k = 0,01$, bestimmt werden:

$$\hat{U}_2 = U_T z, \quad U_2 = \hat{U}_2 / \sqrt{2} \quad (\text{Effektivwert}),$$

$$U_2 = U_T \sqrt{\frac{0,02 \left(\alpha + \gamma \frac{x^2}{8} \right)}{\left(\gamma + \varepsilon \frac{x^2}{8} \right) (2M - 0,01)}}, \quad (46)$$

für $x \ll 1$ gilt

$$U_2 = \sqrt{\frac{\alpha}{\gamma}} \frac{0,1}{\sqrt{M}} U_T. \quad (46')$$

Die Kreuzmodulation M_{k2} mit der 2. Harmonischen $\cos 2pt$ wird

$$M_{k2} = \frac{M}{4} M_k. \quad (47)$$

Die hier gerechneten Ergebnisse sind denjenigen, wie sie für Verzerrungen bei Röhren gefunden wurden [9]—[11], sehr ähnlich.

4. Die Berücksichtigung des Lastwiderstandes

Wird der Transistor am Ausgang nicht mehr kurzgeschlossen, sondern wird ein reeller Lastwiderstand R_l eingeführt, so verändern sich Eingangswiderstand R_{im} und Stromverstärkung $\alpha_{fe m}$. Es gilt [6], [12]:

$$R_{im} = R_b + \frac{U_T \alpha_{fe0 m}}{i_C} \frac{1 + \Phi}{1 + 2\Phi}, \quad (48)$$

$$\text{wobei} \quad \Phi = R_l g_{ce m} \quad (49)$$

das Produkt aus dem Lastwiderstand und dem Momentanwert des im Ersatzschaltbild (Bild 1) auftretenden Leitwertes g_{ce} bezeichnet. Wegen

$$g_{ce} = \frac{I_{OC}}{W} \left. \frac{\partial W}{\partial u_{ce}} \right|_{u=0} \quad (50)$$

$$\text{und} \quad g_{ce m} = \frac{i_C}{W} \left. \frac{\partial W}{\partial u_{ce}} \right|_{u=0}, \quad (50')$$

worin W die Basisschichtdicke darstellt, ist Φ stromabhängig. Der momentane Wert der Stromverstärkung $\alpha_{fe m}$ wird

$$\alpha_{fe m} = \frac{\alpha_{fe0 m}}{1 + 2\Phi}. \quad (51)$$

Es ergibt sich

$$R_{im} = \frac{du}{di_B} = \frac{du}{di_C} \frac{di_C}{di_B} = \frac{du}{di_C} \alpha_{fe m}$$

und mit Gl. (48) und (51)

$$\frac{du}{di_C} = R_b \frac{1 + 2\Phi}{\alpha_{fe0 m}} + \frac{U_T}{i_C} (1 + \Phi). \quad (52)$$

Die Lösung der Differentialgleichung durch Separation der Variablen führt auf

$$u = \int_{I_{OC}}^{i_C} \left[R_b \frac{1 + 2\Phi}{\alpha_{fe0 m}} + \frac{U_T}{i_C} (1 + \Phi) \right] di_C. \quad (53)$$

Die Integration läßt sich nicht einfach durchführen, denn Φ , $\alpha_{fe0 m}$ und $R_{bb'}$ sind stromabhängige Größen. Für Φ wurde der Zusammenhang mit i_C schon gegeben, Gl. (49), (50'). Die Stromverstärkung α_{fe} hängt von Emitterwirkungsgrad, Transportfaktor und Oberflächenrekombination ab [12]—[14] und läßt sich nicht explizit als Funktion des Collectorstromes darstellen.

Eigene Messungen haben gezeigt, daß der Basiszuleitungswiderstand nicht als konstant betrachtet werden darf [15]. In Bild 7 ist für einen Transistor OC 44 der Basiszuleitungswiderstand als Funktion des Collectorgleichstromes aufgetragen. Der Kur-

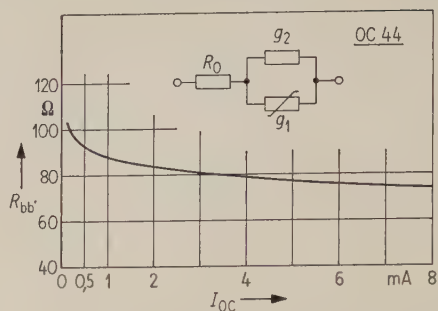


Bild 7. Die Abhängigkeit des Basiszuleitungs-Widerstandes $R_{bb'}$ vom Collectorstrom und das zugehörige Ersatzschema für $R_{bb'}$ mit $R_0 = 68,5 \Omega$, $g_2 = 31 \text{ mA/V}$, $g_1 = 17,25 \cdot I_{OC} \text{ mA/V}$.

venverlauf entspricht nun aber einem Ersatzschaltbild für $R_{bb'}$, wie es ebenfalls in Bild 7 gezeichnet ist. Dabei sind R_0 und g_2 konstante Größen, während g_1 dem Collectorstrom direkt proportional ist. Für diesen speziellen Transistor ergibt sich

$$R_0 = 68,5 \Omega, \quad g_2 = 31 \text{ mA/V}, \quad g_1 = 17,25 \cdot I_{OC} \text{ mA/V},$$

wobei I_{OC} in mA einzusetzen ist.

Wenn nun die Integration von Gl. (53) durchgeführt werden soll, so müssen für R_b , Φ und $\alpha_{fe0 m}$ mittlere, konstante Werte angenommen werden, was gut möglich ist, solange $\Phi \ll 1$ bleibt. Es ergibt sich

$$u = i_C R_b \frac{1 + 2\Phi}{\alpha_{fe0}} + U_T (1 + \Phi) \ln \frac{i_C}{I_{OC}},$$

$$\frac{i_C}{I_{OC}} = \exp \left(\frac{u - \frac{1 + 2\Phi}{\alpha_{fe0}} R_b i_C}{U_T (1 + \Phi)} \right). \quad (54)$$

Nehmen wir nun die Umbezeichnungen

$$R_b (1 + 2\Phi) = R_b^* \quad (55)$$

$$\text{und} \quad U_T (1 + \Phi) = U_T^* \quad (56)$$

vor, so geht Gl. (54) in

$$\frac{i_C}{I_{OC}} = \exp \left(\frac{u - R_b^* \frac{i_C}{\alpha_{fe0}}}{U_T^*} \right) \quad (57)$$

über, und diese Gleichung hat nun genau dieselbe Form wie Gl. (6). An Stelle von q führen wir noch die Abkürzung

$$q^* = \frac{R_b^* I_{OC}}{\alpha_{fe0} U_T^*} \quad \text{bzw.} \quad q^* = \frac{R_b I_{OC}}{\alpha_{fe0} U_T} \frac{1 + 2\Phi}{1 + \Phi} \quad (58)$$

ein und können nun mit den neuen Bezeichnungen alle früher gerechneten Ergebnisse ohne weiteres übernehmen.

Beispielsweise schreibt sich Gl. (30) nun

$$q^* = 0,5, \quad (30')$$

und man ersieht daraus, daß der Arbeitspunkt, für den die 3. Harmonische verschwindet, auch vom Lastwiderstand abhängig ist.

Die Kurvenverläufe der reduzierten Ableitungen bleiben genau gleich, wenn man auf der Abszissenachse q durch q^* ersetzt.

5. Die Verzerrungen über zwei Stufen

Schaltet man zwei Transistorstufen in Kaskade (Bild 8), ohne daß man dazwischen ein Filter einfügt, so ist es interessant zu beobachten, wie sich die Verzerrungen verhalten.

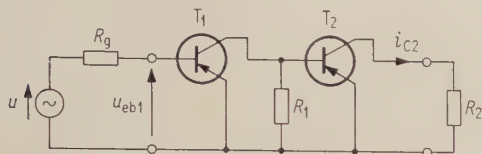


Bild 8. Hochfrequenz-Schema für zweistufigen Verstärker in Emitterschaltung (Speisung fortgelassen).

1.5. Die harmonischen Verzerrungen

Wird am Eingang des ersten Transistors T_1 ein einwelliges Signal $u = \hat{U}_1 \sin \omega_1 t$ angewendet, so sind am Eingang des zweiten (T_2) bereits Harmonische vorhanden, die sich wegen der Nichtlinearität von T_2 mischen. Die Berechnung wird ziemlich umständlich, geht aber prinzipiell, unter Berücksichtigung aller Mischprodukte, genau wie bei der einstufigen Schaltung vor sich. Es sollen hier nur die Ergebnisse angeführt werden. Mit der Abkürzung

$$l = \frac{R_l I_{0C1}}{U_T}, \quad (59)$$

worin R_l den Lastwiderstand von T_1 darstellt, der sich durch Parallelschaltung von R_1 und des Eingangswiderstandes R_{i2m} von T_2 ergibt, erhält man unter der Voraussetzung $R_1 \ll R_{i2m}$ für die Grundharmonische nach dem zweiten Transistor

$$\frac{i_{c2}(\omega = \omega_1)}{I_{0C2}} = -l \left[\alpha_1 \alpha_2 x + (\gamma_1 \alpha_2 + l^2 \alpha_1^3 \gamma_2 + l \alpha_1 \beta_1 \beta_2) \frac{x^3}{8} \right] \sin \omega_1 t. \quad (60)$$

Die Indizes 1 und 2 geben dabei an, um welchen Transistor es sich handelt.

Die zweite Harmonische wird

$$\frac{i_{c2}(\omega = 2\omega_1)}{I_{0C2}} = \pm l(\beta_1 \alpha_2 \mp l \alpha_1^2 \beta_2) \frac{x^2}{4} \cos 2\omega_1 t. \quad (61)$$

Das obere Vorzeichen gilt dabei für die Kaskadierung von zwei p-n-p- oder zwei n-p-n-Transistoren, das untere, wenn ein p-n-p- mit einem n-p-n-Transistor in Kaskade geschaltet wird. Wie man sieht, gelingt es im ersten Fall die zweite Harmonische zum Verschwinden zu bringen. Die Bedingung dafür ist

$$\frac{\beta_1}{\alpha_1^2} = l \frac{\beta_2}{\alpha_2} \quad (62)$$

$$\text{oder} \quad l \frac{1 + q_1^*}{(1 + q_2^*)^2} = 1. \quad (62')$$

In anderen Arbeiten [3] wurde schon gezeigt, wie die zweite Harmonische mittels Push-Pull-Schaltungen zum Verschwinden gebracht werden kann.

Die dritte Harmonische ergibt sich zu

$$\frac{i_{c2}(\omega = 3\omega_1)}{I_{0C2}} = l(\gamma_1 \alpha_2 + l^2 \alpha_1 \gamma_2 + 3l \alpha_1 \beta_1 \beta_2) \frac{x^3}{24} \sin 3\omega_1 t. \quad (63)$$

Da man sowohl γ_1 als auch γ_2 negativ wählen kann, besteht die Möglichkeit, die zweite und die dritte Harmonische gleichzeitig zum Verschwinden zu bringen. Es müssen dann die Bedingungen

$$\gamma_1 \alpha_2 + l^2 \alpha_1 \gamma_2 + 3l \alpha_1 \beta_1 \beta_2 = 0 \quad (64)$$

und

$$\beta_1 \alpha_2 - l \alpha_1^2 \beta_2 = 0$$

gleichzeitig erfüllt sein.

Aus den hier gerechneten Formeln (60), (61) und (63) lassen sich nun, genau wie früher, die Verzerrungsfaktoren bestimmen.

5.2. Die Kreuzmodulation

Die Gl. (44) entsprechende Formel lautet für $x \ll 1$ und $z_p^2 \ll 1$

$$\frac{i_{c2}}{I_{0C2}} = l x \left[\alpha_1 \alpha_2 + \frac{z^2 M}{2} (\alpha_2 \gamma_1 + l \alpha_1 \beta_1 \beta_2) \cdot \left(\cos p t + \frac{M}{4} \cos 2 p t \right) \right] \sin \omega_1 t \quad (65)$$

und die Kreuzmodulation verschwindet für

$$-\frac{\gamma_1}{\alpha_1 \beta_1} = l \frac{\beta_2}{\alpha_2} \quad (66)$$

oder

$$l \frac{1 + q_1^*}{(2q_1^* - 1)(1 + q_2^*)^2} = 1. \quad (66')$$

6. Erweiterung auf höhere Frequenzen

Die in Bild 1 dargestellte Emitterkapazität $C_{b'e}$ besteht im wesentlichen aus der Diffusionskapazität C_D . Letztere ist nun im Bereich kleiner Stromdichten dem Collectorstrom angenähert proportional [5]–[7]:

$$C_D = \frac{I_{0C}}{U_T} \frac{W^2}{2D_p}, \quad (67)$$

wobei W die Basisschichtdicke und D_p die Diffusionskonstante der Löcher im n-Material bezeichnet. (Die Formel würde bei kurzgeschlossenem Collector exakt gelten.)

Für große Eingangssignale ist wieder der Momentanwert des Collectorstromes zu berücksichtigen:

$$C_{Dm} \approx \frac{i_C}{U_T} \frac{W^2}{2D_p}. \quad (67')$$

C_{Dm} bewirkt nun bei höheren Frequenzen einen Abfall der Stromverstärkung, weiter das Auftreten von zusätzlichen Verzerrungen im Collectorstrom und Phasenverschiebungen der einzelnen Anteile des Collectorstromes gegenüber der Eingangsspannung. Dies tritt schon bei Frequenzen auf, bei denen die in Bild 1 eingezeichnete Kapazität $C_{b'e}$, die im wesentlichen die Collectorsperrschicht-Kapazität darstellt, gegenüber C_D weitaus vernachlässigbar ist.

Anschaulich kann man sich das Entstehen der zusätzlichen Verzerrungen leicht erklären, wenn man beachtet, daß C_{Dm} dieselbe Stromabhängigkeit und damit Nichtlinearität besitzt wie $g_{b'e m}$. Die einzelnen, durch C_{Dm} fließenden Stromanteile sind gegenüber den entsprechenden durch $g_{b'e m}$ fließenden um $\pi/2$ phasenverschoben. Diese Ströme erzeugen nun über $R_{bb'}$ Spannungsabfälle, welche ihrerseits wieder den Strom durch $g_{b'e m}$ beeinflussen.

Analog zu Gl. (6) kann auch hier die Übertragungsfunktion gefunden werden:

$$\frac{i_c}{I_{0C}} = \exp \left(\frac{R_b \left(i_c + K \frac{di_c}{dt} \right)}{U_T} \right) \quad (68)$$

mit $K = W^2 \alpha_{fe0} / 2 D_p$.

Aus Gl. (68) bestimmen sich die Verzerrungseffekte, und für die harmonischen Verzerrungen erhält man beispielsweise für die Grundharmonische analog zu Gl. (27)

$$i_c(\omega) = i_c|_{\omega \rightarrow 0} \frac{1 - j \frac{q \omega K}{1 + q}}{1 + \frac{q^2 \omega^2 K^2}{(1 + q)^2}} \quad (69)$$

wobei $i_c|_{\omega \rightarrow 0}$ den in Gl. (27) gerechneten Wert für tiefe Frequenzen bezeichnet. Die Herleitung von Gl. (68), (69) und der übrigen Formeln für die harmonischen Verzerrungen und die Kreuzmodulation sowie die auf sehr ähnliche Weise hergeleitete Theorie für die Mischung soll in weiteren Arbeiten ausführlich behandelt werden.

Symbolliste

a_i	Koeffizienten der Exponentialfunktionen,
A	Stromamplituden,
$C_{b'e}$	Collector-Basis-Kapazität,
$C_{b'e}$	Emitter-Kapazität, $C_{b'e m}$ Momentanwert von $C_{b'e}$,
C_D	Diffusionskapazität,
D_p	Diffusionskonstante der Löcher im n-Material,
e	Betrag der Elementarladung,
f	Frequenz,
$g_{b'e}$	Emitterleitwert, $g_{b'e m}$ Momentanwert von $g_{b'e}$,
g_{ce}	Collectorleitwert, $g_{ce m}$ Momentanwert von g_{ce} ,
i_b	Basisstrom, der von der Eingangswechselspannung herrührt,
i_c	Collectorstrom, der von der Eingangswechselspannung herrührt,
i_C	Momentanwert des Collectorstromes,
I_{0C}	Collector-(Arbeits-) Gleichstrom,
\mathfrak{I}_c	einwelliger Collector-Wechselstrom,
$I_m(x)$	modifizierte Besselsche Funktion m -ter Ordnung mit dem Argument x ,
k	Boltzmannsche Konstante,
k_i	Koeffizienten der Exponentialfunktionen,
$K = W^2 \alpha_{fe0} / 2 D_p$	
$l = R_l I_{0C} / U_T$	
M	Modulationsgrad,

p	Kreisfrequenz des Niederfrequenzsignales,
$q = R_b I_{0C} / U_T \alpha_{fe0}$	
$q^* = R_b^* I_{0C} / U_T^* \alpha_{fe0}$	
$R_b = R_{bb'} + R_g$, $R_b^* = R_b (1 + 2 \Phi)$	
$R_{bb'}$	Basiszuleitungswiderstand,
R_g	Generatorausgangswiderstand,
R_i	Eingangswiderstand,
R_{im}	momentaner Eingangswiderstand,
R_l	Lastwiderstand,
R_z	zusätzlicher Widerstand im Basiskreis,
S	Steilheit,
S_1, S_3, S_5	modifizierte Steilheiten,
S_m	Mischsteilheit,
s	Kreisfrequenz des Störsignales,
t	Zeit,
T	absolute Temperatur,
\hat{U}	einwellige Wechselspannung,
\hat{U}	Wechselspannungsamplitude,
U	effektive Störspannung,
U_T	„thermische Spannung“ $U_T = kT/e$, $U_T^* = U_T (1 + \Phi)$,
$u(t)$	Eingangswechselspannung,
u_{ce}	Ausgangswechselspannung,
W	Basisschichtdicke,
x, y, z	reduzierte Spannungsamplituden $x, y, z = \hat{U} / U_T$,
α_{fe0}	Kurzschluß-Stromverstärkungsfaktor in Emitterschaltung bei tiefen Frequenzen,
α_{fe}	Stromverstärkungsfaktor in Emitterschaltung bei tiefen Frequenzen,
$\alpha_{fe0 m}$ bzw. $\alpha_{fe m}$	Momentanwerte des Stromverstärkungsfaktors,
$\alpha, \beta, \gamma, \delta, \varepsilon, \zeta, \eta, \vartheta$	reduzierte Ableitungen,
Δ	Kreisfrequenz des Pfeiftones,
Φ	reduzierter Lastwiderstand $\Phi = g_{ce} R_l$,
ω	Kreisfrequenz.

Schrifttum

- [1] SPESCHA, G. A. und STRUTT, M. J. O., Theoretische und experimentelle Untersuchungen der Verzerrungen in NF-Flächentransistor-Vierpolen. A.E.Ü. **11** [1957], 307–320.
- [2] ARGÜN, M. und STRUTT, M. J. O., Nichtlineare Verzerrungen einschließlich Kreuzmodulation in HF-Transistorstufen. A.E.Ü. **13** [1959], 227–242.
- [3] MEYER, N. I., Non-linear distortion and small-signal parameters of alloyed junction transistors. Danish Science Press Ltd. Copenhagen 1960.
- [4] LOTSCH, H., Übersicht über die nichtlinearen Verzerrungen einschließlich der Kreuzmodulation in Transistor-Verstärkerstufen. A.E.Ü. **14** [1960], 204–216.
- [5] GIACOLETTO, L. J., Study of p-n-p-alloy-junction transistor from dc through medium frequencies. RCA Rev. **15** [1954], 501–562.
- [6] MOORTGAT-PICK, W., Ersatzschaltbild und Verstärkungseigenschaften des Flächentransistors. A.E.Ü. **13** [1959], 33–48, 82–89.
- [7] VOGEL, J. S., Impedanzmessungen am Transistor mittels Impulsen. Dipl. Aufgabe, Institut für Fernmeldetechnik an der E.T.H. Zürich, Dez. 1958.
- [8] REHWALD, W., Elementare Einführung in die Bessel-, Neumann- und Hankel-Funktionen. S. Hirzel Verlag, Stuttgart 1959.
- [9] STRUTT, M. J. O., Radioempfangsröhren mit großem inneren Widerstand. Hochfrequenztech. u. Elektroakustik **43** [1934], 15–22.
- [10] STRUTT, M. J. O., Elektronenröhren. Springer-Verlag, Berlin 1957.
- [11] STRUTT, M. J. O., Verstärker und Empfänger. Springer-Verlag, Berlin 1951.
- [12] DEWITT, D. und ROSOFF, A. L., Transistor electronics. McGraw-Hill Book Co., New York 1957.
- [13] HAURI, E. R., Zur Frage der Abhängigkeit der Stromverstärkung von Flächentransistoren vom Emitterstrom. Tech. Mitt. PTT (Schweiz), Nr. 11 [1956], 441–451.
- [14] KAUFMANN, P., Der statische Stromverstärkungsfaktor als Funktion des Emitterstromes für Transistoren mit diffundierter und homogener Basisschicht. A.E.Ü. **13** [1959], 141–151.
- [15] GUGGENBUHL, W. und WUNDERLIN, W., Experimentelle Bestimmung des Basis- und Emitterzuleitungswiderstandes von legierten Flächentransistoren mittels Niederfrequenzmessungen. A.E.Ü. **11** [1957], 355–358.

Der Satz von den kleinen Kennlinienänderungen und seine Anwendung auf nichtlineare und lineare Netzwerke

Von EDUARD SCHWARTZ

Mitteilung aus dem Rogowski-Institut für Elektrotechnik der Technischen Hochschule Aachen

(A.E.Ü. 14 [1960], 405–410; eingegangen am 28. Juni 1960)

DK 621.3.052.4

Im 1. Teil dieser Arbeit wird gezeigt, wie man die durch kleine Kennlinienänderungen in nicht-linearen Netzwerken hervorgerufenen Stromänderungen auf einfache und anschauliche Weise mit Hilfe eines Ersatzschaltbildes näherungsweise berechnen kann. Im 2. Teil wird der Sonderfall der linearen Netzwerke ausführlich behandelt.

The first part of this paper shows how, by reference to an equivalent circuit, the current variations caused in nonlinear networks by small variations of the characteristics can be calculated approximately in a simple and illustrative manner. The second part discusses in detail the special case of linear networks.

1. Nichtlineare Netzwerke

1.1. Der Satz von den kleinen Kennlinienänderungen

Betrachtet wird ein Gleichstromnetzwerk mit z Zweigen und $(k + 1)$ Knotenpunkten. Jeder Zweig besteht aus der Reihenschaltung einer Spannungsquelle mit der konstanten Spannung E_n und eines Widerstandes, dessen Spannung U_n eine im allgemeinen nichtlineare Funktion des Zweigstromes I_n ist. Die Bezugspfeile von $U_n(I_n)$ und E_n sollen in jedem Zweig die gleiche Richtung wie die Bezugspfeile von I_n haben; der Index $n = 1, 2, \dots, z$ bezeichnet jeweils die Nummer des betreffenden Zweiges, siehe Bild 1a.

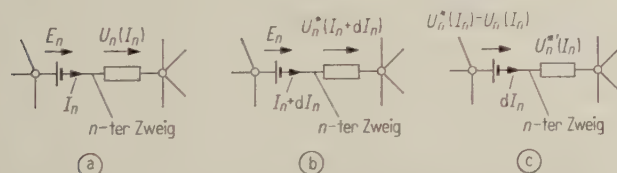


Bild 1. Der Satz von den kleinen Kennlinienänderungen; (a) Ausgangszustand, (b) Zustand nach kleiner Veränderung der Kennlinien, (c) Ersatzschaltbild zur Berechnung der Stromänderungen.

Zur Berechnung der z Zweigströme stehen k unabhängige Knotenpunktgleichungen und $M = z - k$ unabhängige Maschengleichungen zur Verfügung. Diese Gleichungen lassen sich wie folgt allgemein angeben:

Knoten

$$\begin{aligned} 1 \quad & a_{11} I_1 + a_{12} I_2 + \dots + a_{1z} I_z = 0 \\ 2 \quad & a_{21} I_1 + a_{22} I_2 + \dots + a_{2z} I_z = 0 \\ & \vdots \\ k \quad & a_{k1} I_1 + a_{k2} I_2 + \dots + a_{kz} I_z = 0, \end{aligned} \quad (1a)$$

Masche

$$\begin{aligned} 1 \quad & b_{11}[E_1 + U_1(I_1)] + b_{12}[E_2 + U_2(I_2)] + \dots + b_{1z}[E_z + U_z(I_z)] = 0 \\ 2 \quad & b_{21}[E_1 + U_1(I_1)] + b_{22}[E_2 + U_2(I_2)] + \dots + b_{2z}[E_z + U_z(I_z)] = 0 \\ & \vdots \\ M \quad & b_{M1}[E_1 + U_1(I_1)] + b_{M2}[E_2 + U_2(I_2)] + \dots + b_{Mz}[E_z + U_z(I_z)] = 0. \end{aligned} \quad (1b)$$

Hierbei haben die a_{ln} in (1a) folgende Bedeutung: Der 1. Index $l = 1, 2, \dots, k$ bezeichnet den betrachteten Knotenpunkt, der 2. Index $n = 1, 2, \dots, z$ den Zweigstrom. Zählt man die Ströme mit einem zum Knotenpunkt gerichteten Bezugspfeil als positiv, dann gilt

$$a_{ln} = \begin{cases} +1, & \text{wenn der Bezugspfeil von } I_n \text{ auf den } l\text{-ten Knotenpunkt hinweist,} \\ 0, & \text{wenn } I_n \text{ am } l\text{-ten Knotenpunkt nicht vorhanden ist,} \\ -1, & \text{wenn der Bezugspfeil von } I_n \text{ vom } l\text{-ten Knotenpunkt wegweist.} \end{cases}$$

In (1b) bezeichnet der 1. Index $m = 1, 2, \dots, M$ der b_{mn} die betreffende Masche (deren Umlaufsinn willkürlich gewählt wird), der 2. Index $n = 1, 2, \dots, z$ den Zweig mit der Gesamtspannung $E_n + U_n(I_n)$ entsprechend Bild 1a. Es ist

$$b_{mn} = \begin{cases} +1, & \text{wenn der Umlaufsinn der } m\text{-ten Masche die gleiche Richtung wie der Bezugspfeil von } I_n \text{ hat,} \\ 0, & \text{wenn die } m\text{-te Masche den } n\text{-ten Zweig nicht enthält,} \\ -1, & \text{wenn der Umlaufsinn der } m\text{-ten Masche die entgegengesetzte Richtung wie der Bezugspfeil von } I_n \text{ hat.} \end{cases}$$

Nunmehr wird ein von dem durch das Gleichungssystem (1) beschriebenen nur wenig abweichender Zustand betrachtet: Die Kennlinien der nicht-linearen Widerstände sollen verändert werden, siehe Bild 2. Als Beispiele für Veränderungen von Kennlinien seien angeführt: Erwärmung der Widerstände infolge der sie durchfließenden Ströme (innere Ursache), Auswechseln von Widerständen, z. B. Röhren, mit unvermeidbaren Kennlinienstreuungen (äußere Ursache). Hierbei soll aber vorausgesetzt

werden, daß zumindest in der Umgebung der in Frage kommenden Arbeitspunkte die ursprünglichen und die neuen Kennlinien nur wenig voneinander abweichen (was bei den genannten Beispielen der Fall ist) und daß alle Arbeitspunkte stabil sind, so daß auch die sich jetzt einstellenden Ströme im Netzwerk nur wenig von den ursprünglichen verschieden sind: $I_n + dI_n$. Dies bedeutet eine Arbeitspunktverschiebung von A nach B in Bild 2.

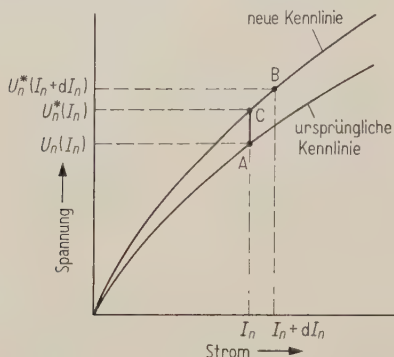


Bild 2. Kleine Veränderung der Kennlinien.

Die ursprünglichen Spannungen $U_n(I_n)$ an den nichtlinearen Widerständen werden dadurch um zwei Anteile verändert. Der erste wird durch die Kennlinienänderung (Arbeitspunktverschiebung von A nach C) und der zweite durch die Stromänderung (Arbeitspunktverschiebung von C nach B) hervorgerufen. Mit den Bezeichnungen von Bild 2 werden die neuen Spannungen an den nichtlinearen Widerständen in 1. Näherung

$$U_n^*(I_n + dI_n) = U_n^*(I_n) + U_n'(I_n) dI_n \quad (2)$$

mit

$$U_n' = dU_n^*/dI_n.$$

Hierbei wird durch den Stern angedeutet, daß es sich um die neuen, geänderten Kennlinien handelt.

Für diesen, durch die kleinen Kennlinienänderungen hervorgerufenen, neuen Zustand des Netzwerkes (angedeutet in Bild 1b) gilt das Gleichungssystem (1a) mit $I_n + dI_n$ statt I_n und (1b) mit $U_n^*(I_n + dI_n)$ nach (2) statt $U_n(I_n)$. Subtrahiert man von den Gleichungen des so entstehenden Systems die entsprechenden des ursprünglichen, so erhält man

$$\begin{aligned} a_{11}dI_1 + a_{12}dI_2 + \dots + a_{1z}dI_z &= 0 \\ a_{21}dI_1 + a_{22}dI_2 + \dots + a_{2z}dI_z &= 0 \\ \vdots & \\ a_{k1}dI_1 + a_{k2}dI_2 + \dots + a_{kz}dI_z &= 0, \end{aligned} \quad (3a)$$

$$\begin{aligned} b_{11}[U_1^*(I_1) - U_1(I_1) + U_1'(I_1)dI_1] + b_{12}[U_2^*(I_2) - U_2(I_2) + U_2'(I_2)dI_2] + \\ + \dots + b_{1z}[U_z^*(I_z) - U_z(I_z) + U_z'(I_z)dI_z] &= 0 \\ b_{21}[U_1^*(I_1) - U_1(I_1) + U_1'(I_1)dI_1] + b_{22}[U_2^*(I_2) - U_2(I_2) + U_2'(I_2)dI_2] + \\ + \dots + b_{2z}[U_z^*(I_z) - U_z(I_z) + U_z'(I_z)dI_z] &= 0 \\ \vdots & \\ b_{M1}[U_1^*(I_1) - U_1(I_1) + U_1'(I_1)dI_1] + b_{M2}[U_2^*(I_2) - U_2(I_2) + U_2'(I_2)dI_2] + \\ + \dots + b_{Mz}[U_z^*(I_z) - U_z(I_z) + U_z'(I_z)dI_z] &= 0. \end{aligned} \quad (3b)$$

Vergleicht man die Gleichungssysteme (3) und (1), so erkennt man, daß (3) aus (1) hervorgeht, wenn man in (1a) alle I_n durch dI_n , in (1b) alle E_n durch $U_n^*(I_n) - U_n(I_n)$ und alle $U_n(I_n)$ durch $U_n'(I_n) dI_n$ ersetzt. Man kann (3) also deuten als Gleichungssystem eines Netzwerkes, das dem ursprünglichen gleich ist, bei dem aber

1. statt der ursprünglichen Spannungsquellen, die als kurzgeschlossen zu betrachten sind, solche mit der Spannung $U_n^*(I_n) - U_n(I_n)$ (das ist die Spannungsänderung zwischen den Punkten A und C in Bild 2) wirksam sind,
2. statt der ursprünglichen Ströme die Ströme dI_n fließen,
3. die ursprünglichen, nichtlinearen Widerstände mit der Spannung $U_n(I_n)$ durch Zweipole mit der Spannung $U_n'(I_n) dI_n$ ersetzt sind (diese Spannung entspricht der Spannungsänderung von C nach B in Bild 2). Entsprechend dem Ohmschen Gesetz kann man diese Zweipole auffassen als Widerstände $U_n'(I_n)$, die von den Strömen dI_n durchflossen werden. Diese Widerstände $U_n'(I_n)$ sind als linear anzusehen, da sie von den jetzt fließenden Strömen dI_n nicht abhängen.

Aus diesen Überlegungen folgt der

Satz 1 („Satz von den kleinen Kennlinienänderungen“):

Werden in einem Gleichstromnetzwerk, ausgehend von einem bestimmten, stabilen Zustand (Bild 1a), die Kennlinien der nichtlinearen Widerstände ein wenig geändert (Bild 1b), so lassen sich die dadurch hervorgerufenen Stromänderungen näherungsweise mit Hilfe eines Netzwerkes gleicher Struktur wie das ursprüngliche berechnen, in dem die ursprünglichen Spannungsquellen kurzgeschlossen sind, dafür ist in jedem Zweig eine Spannungsquelle mit der Spannung $U_n^(I_n) - U_n(I_n)$ und statt des ursprünglichen nichtlinearen Widerstandes ein linearer von der Größe $U_n'(I_n)$ wirksam (Bild 1c).*

Hierbei sind $U_n^*(I_n)$ bzw. $U_n(I_n)$ die Spannungen auf den neuen bzw. alten Kennlinien an der Stelle der ursprünglichen Ströme (vgl. Bild 2), $U_n'(I_n)$ die Ableitungen der neuen Kennlinien (differentielle Widerstände) an der Stelle der ursprünglichen Ströme, wobei häufig ohne nennenswerten Verlust an Genauigkeit auch die Ableitungen der ursprünglichen Kennlinien an der Stelle der alten Ströme eingesetzt werden können.

Bemerkenswert an dem durch Satz 1 ausgedrückten „Ersatzschaltbild“ zur Berechnung der Stromänderungen sind vor allem folgende Punkte:

1. Bei gegebenem Ausgangszustand und gegebenen ursprünglichen und veränderten Kennlinien sind die Spannungsquellen und Widerstände des Ersatzschaltbildes nur abhängig von den Strömen des Ausgangszustandes, also bekannt.
2. Das Ersatzschaltbild ist ein lineares Netzwerk, da seine Widerstände $U_n^*(I_n)$ zwar von den ursprünglichen, nicht aber von den im Ersatzschaltbild fließenden Strömen dI_n abhängen.

Aus der Linearität des Ersatzschaltbildes folgt der Satz 2 („Überlagerungsgesetz für kleine Stromänderungen in nichtlinearen Netzwerken“):

Die durch mehrere, gleichzeitig an verschiedenen nichtlinearen Widerständen durchgeführte, kleine Kennlinienänderungen in stabilen Gleichstromnetzwerken hervorgerufenen Stromänderungen können näherungsweise für jede Kennlinienänderung einzeln berechnet und dann linear überlagert werden.

Außer dem Überlagerungsgesetz können natürlich auf das Ersatzschaltbild alle anderen Berechnungsmethoden linearer Netzwerke angewendet werden. Dadurch ist die Berechnung kleiner Abweichungen von einem bekannten Ausgangszustand infolge von Kennlinienänderungen leicht durchführbar, während die Berechnung des Ausgangszustandes bei komplizierten nichtlinearen Netzwerken in der Regel explizit nicht möglich ist und umständliche numerische Arbeit erfordert.

1.2. Beispiele

1.2.1. Reihenschaltung zweier nichtlinearer Widerstände

An einer Gleichspannung $E = 400 \text{ V}$ liegen in Reihe zwei nichtlineare Widerstände mit den Kennlinien

$$U_1 = a I^2, \quad U_2 = b I^2 + c I$$

(gültig im 1. Quadranten). Es sei z. B.

$$a = 60 \text{ V/A}^2, \quad b = 40 \text{ V/A}^2 \quad \text{und} \quad c = 300 \text{ V/A}.$$

Dann fließt, wie man leicht bestätigt, durch beide Widerstände ein Strom $I_a = 1 \text{ A}$. Wie groß ist die Änderung dieses Stromes dI_a , wenn die Kennlinien die nur wenig abweichende Form

$$U_1^* = a^* I^2, \quad U_2^* = b^* I^2 + c^* I$$

mit

$$a^* = 61 \text{ V/A}^2, \quad b^* = 41 \text{ V/A}^2 \quad \text{und} \quad c^* = 296 \text{ V/A}$$

haben? Die für das Ersatzschaltbild notwendigen Größen sind die (differentiellen) Widerstände

$$U_1^{*'}(I_a) = 2a^* I_a = 122 \Omega,$$

$$U_2^{*'}(I_a) = 2b^* I_a + c^* = 378 \Omega$$

und die Spannungsquellen

$$U_1^*(I_a) - U_1(I_a) = a^* I_a^2 - a I_a^2 = 1 \text{ V},$$

$$U_2^*(I_a) - U_2(I_a) =$$

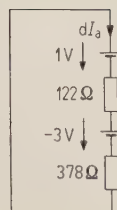
$$= b^* I_a^2 + c^* I_a - b I_a^2 - c I_a = -3 \text{ V}.$$

Hiermit erhält man nach Satz 1 Bild 3 zur Berechnung von dI_a . Mit Hilfe dieses Bildes findet man für dI_a sofort

$$dI_a = +2 \text{ V} / 500 \Omega = +4 \text{ mA}.$$

Die bei diesem einfachen Beispiel exakt durchführbare Rechnung liefert den Wert $+3,997 \text{ mA}$!

Bild 3. Ersatzschaltbild zur Berechnung der Stromänderung bei der Reihenschaltung zweier nichtlinearer Widerstände.



1.2.2. Anwendung auf verzweigte magnetische Kreise

Bei verzweigten magnetischen Kreisen, deren Flüsse im wesentlichen durch Eisenschenkel verlaufen, läßt sich Satz 1 auch anwenden, und zwar zur Berechnung von Flußänderungen infolge kleiner Kennlinienänderungen. Hierbei sind die Spannungsquellen E_n durch die elektrischen Durchflutungen Θ_n , die Ströme I_n durch die Flüsse Φ_n und die Funktionen $U_n(I_n)$ durch die magnetischen Kennlinien $V_n(\Phi_n)$, siehe Bild 4, wobei $V_n = H_n l_n$ die magnetischen Spannungen sind, zu ersetzen. Wegen der Einfachheit dieser bekannten Analogie sind weitere Ausführungen nicht notwendig.

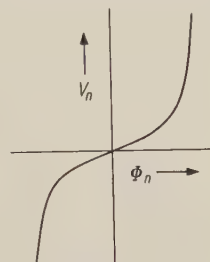


Bild 4. Magnetische Kennlinie.

2. Lineare Netzwerke

2.1. Der Satz von den kleinen Widerstandsänderungen

Selbstverständlich gilt Satz 1 auch für lineare Netzwerke, bei denen die Spannung-Strom-Kennlinien Geraden sind. Dieser Sonderfall soll wegen der außerordentlichen Bedeutung der linearen Netzwerke getrennt behandelt werden. Es wird sich zeigen, daß mit Hilfe der in diesem Fall entstehenden Formulierung des Satzes 1 nicht nur die Stromänderungen infolge kleiner Widerstandsänderungen (Beispiel 2.2.1) berechnet, sondern auch netzwerktheoretische Fragen (Beispiele 2.2.2 und 2.2.3) behandelt werden können.

Im Fall linearer Widerstände gilt $U_n(I_n) = R_n I_n$, wobei R_n der stromunabhängige Widerstand des n -ten Zweiges ist. Eine kleine Kennlinienänderung bedeutet in diesem Fall eine kleine Änderung des Widerstandes: $R_n + dR_n$. Damit wird

$$U_n^*(I_n) - U_n(I_n) = (R_n + dR_n) I_n - R_n I_n = I_n dR_n, \\ U_n^*(I_n) = R_n + dR_n \approx R_n \quad \text{wegen} \quad dR_n \ll R_n. \quad (4)$$

Aus Satz 1 erhält man hiermit den

Satz 3 („Satz von den kleinen Widerstandsänderungen“):

Werden in einem linearen Gleichstromnetzwerk, ausgehend von einem bestimmten Zustand (Bild 5a), die Widerstände ein wenig geändert (Bild 5b), so lassen sich die dadurch hervorgerufenen Stromänderungen näherungsweise mit Hilfe eines Netzwerkes gleicher Struktur wie das ursprüngliche berechnen, in dem die ursprünglichen Spannungsquellen kurzgeschlossen

sind, dafür ist in jedem Zweig eine Spannungsquelle $I_n dR_n$ wirksam, während die Widerstände die gleichen wie im Ausgangszustand bleiben (Bild 5c).

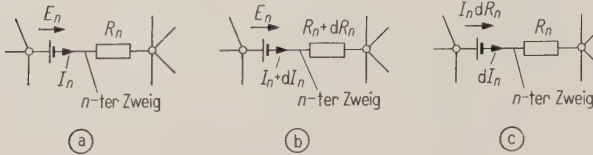


Bild 5. Der Satz von den kleinen Widerstandsänderungen; (a) Ausgangszustand, (b) Zustand nach kleiner Veränderung der Widerstände, (c) Ersatzschaltbild zur Berechnung der Stromänderungen.

Hierzu noch einige Bemerkungen:

1. Satz 3 ist, für nur eine Widerstandsänderung formuliert, als „Compensation Theorem“ bekannt [1], [2].
2. Das Ersatzschaltbild gilt um so genauer, je kleiner die dR_n gegenüber den R_n sind. Es gilt exakt bei unendlich kleinen Widerstandsänderungen im Sinne der Differentialrechnung.
3. Auf das Ersatzschaltbild zur Berechnung der Stromänderungen (angedeutet in Bild 5c) kann wegen dessen Linearität das Überlagerungsgesetz angewendet werden, was im vorliegenden Fall linearer Ausgangsnetzwerke einleuchtender erscheint als bei nichtlinearen (Satz 2).
4. Satz 3 ist auch für lineare Wechselstromnetzwerke gültig. Hierbei ist \underline{E}_n statt E_n , \underline{I}_n statt I_n und \underline{R}_n statt R_n zu setzen.
5. Vernachlässigt man in Gleichung (4) dR_n nicht, so erhält man auch bei beliebig großen Widerstandsänderungen ΔR_n ein exakt gültiges Ersatzschaltbild zur Berechnung der Stromänderungen. (Grund: Gl. (2), die einzige Näherung bei der Ableitung des Satzes 1, gilt bei Geraden exakt.) Im Ersatzschaltbild sind dann die Span-

nungsquellen $I_n \Delta R_n$ und die Widerstände $R_n + \Delta R_n$ wirksam. Diese Formulierung des Satzes 3 kann gegebenenfalls (z. B. bei Rechnungen mit hohen Anforderungen an die Genauigkeit) von Nutzen sein.

2.2. Beispiele

2.2.1. Wheatstone-Brücke

Die in Bild 6a dargestellte Wheatstone-Brücke sei abgeglichen:

$$R_1/R_2 = R_3/R_4, \quad I_5 = 0, \quad I_1 = E/(R_1 + R_2).$$

Der zu messende Widerstand R_1 ändere sich um dR_1 , Bild 6b. Welchen Strom dI_5 durch den „Galvanometerzweig“ hat diese Änderung zur Folge?

Nach Satz 3 findet man zur Berechnung von dI_5 das Ersatzschaltbild 6c. Auf Grund dieses Ersatzschaltbildes erhält man

$$dI_5 = - \frac{R_2 E dR_1}{(R_1 + R_2)^2 \left(R_5 + \frac{R_1 R_2}{R_1 + R_2} + \frac{R_3 R_4}{R_3 + R_4} \right)}.$$

Diesen Ausdruck kann man zur Berechnung der Empfindlichkeit der Brücke benutzen (siehe z. B. [3]). Bemerkenswert ist aber hier nicht das Ergebnis selbst, sondern daß es ohne Durchführung der Differentiation dI_5/dR_1 mit Hilfe eines anschaulichen Ersatzschaltbildes gewonnen wurde.

2.2.2. Ersatzzweipolquelle

In einem beliebigen, aber linearen Netzwerk (Bild 7a) fließt der Strom I_a durch den Widerstand R_a . Der Innenwiderstand R_i des Netzwerkes gemäß Bild 7b sei bekannt. Bei einer differentiellen Änderung des Widerstandes R_a erhält man nach Satz 3 Bild 7c zur Berechnung der Stromänderung dI_a . Hieraus liest man ab

$$dI_a = - \frac{I_a dR_a}{R_a + R_i} \quad \text{oder} \quad \frac{dI_a}{I_a} = - \frac{dR_a}{R_a + R_i}.$$

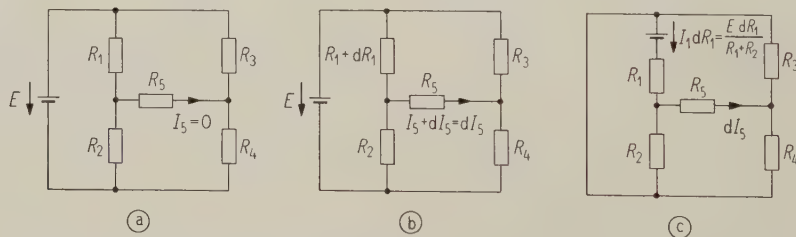


Bild 6. Wheatstone-Brücke; (a) abgeglichener Zustand, (b) kleine Abweichung vom Abgleich, (c) Ersatzschaltbild zur Berechnung von dI_5 .

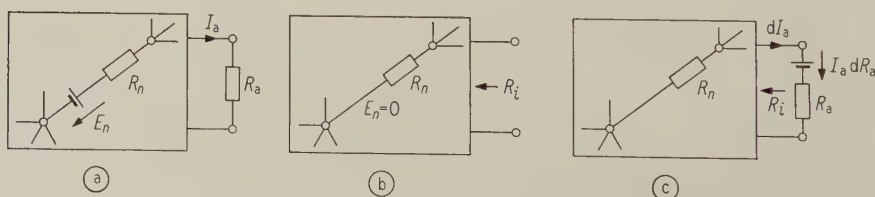


Bild 7. (a) Beliebige lineares Netzwerk, (b) Innenwiderstand desselben (Kurzschluß aller Spannungsquellen E_n), (c) Ersatzschaltbild zur Berechnung von dI_a .

Diese Differentialgleichung für I_a als Funktion von R_a läßt sich leicht integrieren. Die einander auf der rechten und linken Seite entsprechenden Grenzen sind:

$$\begin{aligned} R_a &\rightarrow I_a, \text{ beliebiger Zustand,} \\ 0 &\rightarrow I_{ak}, \text{ Kurzschluß.} \end{aligned}$$

Hiermit wird

$$\begin{aligned} \int_{I_{ak}}^{I_a} \frac{dI_a}{I_a} &= - \int_0^{R_a} \frac{dR_a}{R_a + R_1} = \int_{R_a}^0 \frac{dR_a}{R_a + R_1}, \\ \ln \frac{I_a}{I_{ak}} &= \ln \frac{R_1}{R_a + R_1}, \\ I_a &= \frac{R_1 I_{ak}}{R_1 + R_a}. \end{aligned}$$

Dies ist der Satz von der Ersatzzweipolquelle (siehe z. B. [4]). Auch hier ist nicht das Ergebnis selbst interessant, sondern die vermutlich neue Art der Ableitung.

2.2.3. Der Satz von der Eingangswiderstandsänderung und seine Anwendung auf Reaktanzzweipole

An einem linearen Zweipol, der nur aus Widerständen besteht (also keinerlei Spannungsquellen enthält), liegt die konstante Spannung E , der von dem Zweipol aufgenommene Strom ist I , siehe Bild 8a. Der Zweipolwiderstand ist dann

$$R = E/I.$$

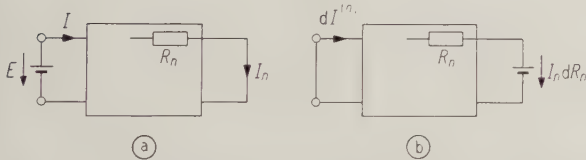


Bild 8. (a) Ausgangszustand eines Zweipoles, (b) Ersatzschaltbild für den Fall einer Zweigwiderstandsänderung.

Ändern sich in dem Zweipol alle Zweigwiderstände R_n um dR_n , so ändert sich auch der Strom I um dI und damit der Zweipolwiderstand:

$$\begin{aligned} R + dR &= E/(I + dI) = R - R dI/I, \\ dR &= -R dI/I = -E dI/I^2. \end{aligned} \quad (5)$$

Ändert sich in dem Zweipol nur ein Zweigwiderstand R_n um dR_n , dann läßt sich die zugehörige Änderung $dI^{(n)}$ des Zweipolstromes I mit Hilfe des Ersatzschaltbildes 8b berechnen, das man nach Satz 3 angeben kann. Durch Vergleich mit dem Ausgangszustand Bild 8a erhält man nach dem Umkehrungssatz (siehe z. B. [5])

$$\begin{aligned} E/I_n &= -I_n dR_n/dI^{(n)}, \\ dI^{(n)} &= -I_n^2 dR_n/E. \end{aligned}$$

Ändern sich mehrere oder alle Zweigwiderstände des Zweipoles, so dürfen die $dI^{(n)}$ überlagert werden (vergleiche Bemerkung 3 zu Satz 3). Die gesamte Änderung des Zweipolstromes wird mithin

$$dI = \sum_1^z dI^{(n)} = -\frac{1}{E} \sum_1^z I_n^2 dR_n.$$

Setzt man dI in die oben gefundene Gleichung (5) für dR ein, so erhält man den

Satz 4 („Satz von der Eingangswiderstandsänderung linearer, nur aus Widerständen bestehender Netzwerke“):

Verändert man in einem linearen Netzwerk, das nur aus Widerständen besteht, die Zweigwiderstände R_n um dR_n , dann verändert sich der zwischen zwei beliebigen Klemmen des Netzwerkes gemessene Eingangswiderstand um

$$dR = + \sum_1^z \left(\frac{I_n}{I} \right)^2 dR_n.$$

Diese Gleichung für die Änderung des Eingangswiderstandes enthält die einleuchtende Aussage, daß Widerstandsänderungen in Zweigen mit kleinen Strömen den Eingangswiderstand nur wenig, in Zweigen mit großen Strömen dagegen stark beeinflussen; Widerstandsänderungen in stromlosen Zweigen haben keinen Einfluß auf den Eingangswiderstand.

Dieser Satz von der Eingangswiderstandsänderung läßt sich auch mit der in [6] abgeleiteten Beziehung

$$\frac{\partial R}{\partial R_n} = \left(\frac{I_n}{I} \right)^2 \quad (6)$$

unter Zuhilfenahme des totalen Differentials beweisen. Umgekehrt kann man Gl. (6) als Sonderfall von Satz 4 für nur eine Widerstandsänderung im Netzwerk betrachten. Nach dieser Gleichung sind die Vorzeichen von ∂R und ∂R_n gleich in einem Netzwerk aus nur ohmschen Widerständen. Hieraus folgt für diesen Sonderfall folgende anschauliche Formulierung des Satzes von der Eingangswiderstandsänderung: Vergrößert (verkleinert) man in einem linearen Netzwerk, das nur aus ohmschen Widerständen besteht, irgendeinen Zweigwiderstand R_n um ∂R_n , dann wird auch der zwischen zwei beliebigen Klemmen des Netzwerkes gemessene Eingangswiderstand größer (kleiner) oder bleibt konstant, falls $I_n = 0$ ist.

Wie in der Bemerkung 4 zu Satz 3 gesagt wurde, dürfen die abgeleiteten Ergebnisse auch auf Wechselstromnetzwerke angewendet werden. In diesem Fall wird die Eingangswiderstandsänderung

$$d\underline{Z} = + \sum_1^z \left(\frac{\underline{I}_n}{\underline{I}} \right)^2 d\underline{Z}_n.$$

Diese Gleichung soll auf Reaktanznetzwerke (die nur aus verlustlosen Spulen und Kondensatoren bestehen) angewendet werden. Hierbei ist

$$\begin{aligned} \underline{Z} &= jX, \\ \underline{Z}_n &= j(\omega L_n - 1/\omega C_n). \end{aligned}$$

Denkt man sich die Frequenz differentiell um $d\omega$ verändert, so ändern sich auch die Zweigwiderstände:

$$\underline{Z}_n + d\underline{Z}_n = j[(\omega + d\omega)L_n - 1/(\omega + d\omega)C_n],$$

$$d\underline{Z}_n = j(L_n + 1/\omega^2 C_n)d\omega.$$

Damit wird

$$d\underline{Z} = j dX = j d\omega \sum_1^z \left(\frac{I_n}{I} \right)^2 (L_n + 1/\omega^2 C_n).$$

Bedenkt man, daß bei Reaktanzzweipolen alle Zweigströme \underline{I}_n entweder in Phase oder in Gegenphase mit dem Zweipolstrom \underline{I} sind, so erkennt man, daß

$$dX/d\omega > 0 \quad (7)$$

gilt, eine bekannte Gesetzmäßigkeit bei Reaktanzzweipolen (siehe z. B. [7]).

Setzt man in Satz 4

$$R = 1/G, R_n = 1/G_n, I = GE \text{ und } I_n = G_n U_n,$$

so erhält man den dualen Satz, der die Änderung des Eingangsleitwertes linearer, nur aus Widerständen bestehender Netzwerke betrifft:

$$dG = + \sum_1^z \left(\frac{U_n}{E} \right)^2 dG_n.$$

Mit dieser Gleichung läßt sich z. B. die der Beziehung (7) entsprechende für den Leitwert von Reaktanzzweipolen ableiten.

Zum Schluß dieser Arbeit möchte der Verfasser Herrn Dr.-Ing. H.-J. BUTTERWECK für nützliche Diskussionen und Hinweise danken.

Schrifttum

- [1] SHEA, T. E., Transmission networks and wave filters; 11. Aufl. Van Nostrand Co., New York 1958, S. 56–58.
- [2] MONTGOMERY, C. G., DICKE, R. H. und PURCELL, E. M., Principles of microwave circuits. McGraw-Hill Book Co., New York 1948, S. 92–93.
- [3] SCHWERTFEGGER, W., Elektrische Meßtechnik II. C. F. Winter'sche Verlagshandlung, Leipzig 1939, S. 75–82.
- [4] KÜPFMÜLLER, K., Einführung in die theoretische Elektrotechnik. 5. Aufl. Springer-Verlag, Berlin 1955, S. 22.
- [5] Wie [4], S. 322–325.
- [6] VRATSANOS, J., Zur Berechnung der Stromverteilung in einem linearen Netzwerk. A.E.Ü. II [1957], 76–80.
- [7] CAUER W., Theorie der linearen Wechselstromschaltungen, 2. Aufl. Akademie-Verlag, Berlin 1954, S. 181.

BUCHBESPRECHUNGEN

W. A. Geyger, Magnetverstärker-Schaltungen ; Grundlagen, Eigenschaften, Anwendungen. Verlag Berliner Union, Stuttgart 1959, 374 Seiten, 183 Bilder, 15,5 cm × 23,0 cm, Ganzleinen DM 64,—.

Magnetverstärkerschaltungen haben in den letzten Jahren als Bestandteile elektronischer Einrichtungen außerordentlich an Bedeutung zugenommen. So wurde eine große Zahl von Schaltungen für die verschiedensten Aufgaben mit speziellen Eigenschaften entwickelt. Der Verfasser dieses Buches ist seit vielen Jahren führend auf dem Gebiet der Magnetverstärker tätig. So konnte er in hervorragender Weise Ordnung in die Vielfalt der Schaltungsmöglichkeiten bringen.

Nach einer Einleitung mit einem interessanten historischen Überblick und Kennenlernen der Bestandteile des Magnetverstärkers werden in den Kapiteln 3 bis 17 die verschiedenen Schaltungsarten erläutert und mit vielen Schaltbeispielen (teilweise mit Dimensionierungsangaben) belegt. Zum Schluß werden statische und dynamische Eigenschaften sowie Anwendungen von Magnetverstärkern behandelt.

Mit einfachen und anschaulichen Erklärungen werden die wesentlichen Eigenschaften der Schaltungen aufgezeigt. Theoretische Abhandlungen enthält das Buch nicht. Hierfür wird in umfassenden Literaturverzeichnissen, die jedem Kapitel angehängt sind, auf die Originalarbeit verwiesen.

So ist das Buch sowohl für den Lernenden wie für den Ingenieur, der elektronische Schaltungen zu entwickeln hat, eine gute Einführung sowie ein umfassendes, reichhaltiges Nachschlagewerk.

Wir sind dankbar, daß die zweite, auf den neuesten Stand gebrachte Auflage dieses Buches nunmehr auch in deutscher Sprache vorliegt.

R. ELSNER

L. A. Chegny, Wave propagation in a random medium. (Aus dem Russischen übersetzt von R. A. SILVERMANN.) McGraw-Hill Book Co., New York-London-Toronto 1960, VIII, 168 Seiten, 13 Bilder, 19 cm × 25,5 cm, Ganzleinen sh 58/—.

In den vergangenen anderthalb Jahrzehnten wurden die Einflüsse kleiner Inhomogenitäten des Ausbreitungsmediums, die statistisch verteilt sind, auf elektrische und akustische Wellen von einer größeren Anzahl von Forschern in der ganzen Welt eingehend studiert. Es liegt hier ein von einem

bedeutenden sowjetischen Forscher verfaßtes Buch vor (Mitglied der Akademie der Wissenschaften der Sowjet-Union), das von einem amerikanischen Fachmann ins Englische übersetzt ist, der selbst auf dem genannten Gebiet wohl bekannt ist.

Das Buch geht aus von der Störung skalarer (Schall-) Wellen durch kleine Inhomogenitäten im Ausbreitungsmedium. Es soll hier ein Überblick über die wesentlichen Abschnitte nach dem Inhaltsverzeichnis gegeben werden:

Teil I: Strahlentheorie

- Kap. 1. Statistische Eigenschaften des Mediums
- Kap. 2. Strahlenstatistik

Teil II: Beugungstheorie

- Kap. 3. Die Wellengleichung
- Kap. 4. Streuung durch Inhomogenitäten
- Kap. 5. Fluktuationen
- Kap. 6. Korrelationen der Fluktuationen

Teil III: Der Einfluß der Fluktuationen auf das Beugungsbild eines fokussierenden Systems

- Kap. 7. Allgemeine Formeln
- Kap. 8. Die mittlere Verteilung in der Nähe des Fokus
- Kap. 9. Fluktuationen hinter der Linse

2 Anhänge, Literaturverzeichnis.

Wie man aus diesem Auszug aus dem Inhaltsverzeichnis sieht, ist der genannte Problemkreis sehr systematisch und umfassend in der für Wellenprobleme sinnvollen Reihenfolge behandelt; zuerst die „Strahlenoptik“, dann die Wellenstreuung mit ihren zeitabhängigen Schwankungen.

Man erhält Kenntnis von einer großen Zahl von russischen Arbeiten, deren Ergebnisse nun in englischer Sprache vorliegen; auch sind viele amerikanische und englische Arbeiten berücksichtigt.

Dabei ist auch ein bisher im Westen nur wenig studiertes Gebiet, nämlich der Einfluß von Fluktuationen des Brechungsvermögens auf Linsensysteme, sehr ausführlich behandelt.

Die Darstellung ist klar und verständlich, einige komplizierte Integrale sind im Anhang gesondert behandelt. Das Literaturverzeichnis ist sehr umfangreich.

Das Buch kann jedem, der sich mit der Wellenstreuung an fluktuierenden Inhomogenitäten des Mediums zu befassen hat, bestens empfohlen werden.

G. ECKART

Die Steilheit des HF-Legierungs- und Drifttransistors als Funktion der Frequenz und des Arbeitspunktes und ihre Ableitungen

Von WILLY MINNER

Mitteilung aus dem Röhrenlaboratorium der Telefunken G.m.b.H., Ulm (Donau)

(A.E.Ü. 14 [1960], 411—420; eingegangen am 1. Juli 1960)

DK 621.382.333.3

Die frequenzabhängige HF-Steilheit eines Transistors läßt sich nicht aus der Collectorgleichstrom-Basisemittergleichspannungs-Kennlinie herleiten. In dieser Arbeit wird eine empirisch ermittelte Gleichung angegeben, die die HF-Steilheit als Funktion der Basisemitter-Gleichspannung U_{BE} bis zu etwa der doppelten α -cutoff-Frequenz zu berechnen gestattet, wenn die Steilheit als Funktion von U_{BE} bei einer Frequenz, die in der Nähe der α -cutoff-Frequenz liegt, aus Messungen vorliegt. Durch Differenzieren dieser Gleichung nach U_{BE} erhält man die höheren Ableitungen der Steilheit, die zur Berechnung der Mischsteilheit und des Collectorkurzschlußstromes mit Kombinationsfrequenzen höherer Ordnung benötigt werden. In der angegebenen Grundgleichung und den davon abgeleiteten Gleichungen wird die Zeitkonstante des HF-Flächentransistors unterteilt in die

für die Basis-Emitterstrecke wirksame Zeitkonstante τ_B
und die für die Collector-Basisstrecke gültige Zeitkonstante τ_C .

Für die Berechnung der Mischsteilheit ist die Zeitkonstante τ_B mit der Hochfrequenz (Eingangsfrequenz) und die Zeitkonstante τ_C mit der Zwischenfrequenz (Ausgangsfrequenz) wirksam. Durch vergleichende Messungen der Geradeaussteilheit und der Mischsteilheit ist es somit möglich, die beiden Zeitkonstanten eines HF-Transistors zu trennen. Wie die Meßergebnisse zeigen, ist τ_C für den HF-Legierungstransistor klein gegen τ_B , so daß die Mischsteilheit nur wenig von der gewählten Zwischenfrequenz abhängt. Anders verhält es sich beim Drifttransistor; hier ist τ_C sogar größer als τ_B , so daß eine starke Abhängigkeit der Mischsteilheit von der gewählten Zwischenfrequenz vorhanden ist.

Die in vorliegender Arbeit angegebenen Gleichungen, die von der empirisch gefundenen Grundgleichung abgeleitet sind, wurden für den HF-Legierungstransistor AF 101 und den Drifttransistor OC 615 experimentell überprüft. Für die behandelten Fälle, Betrag der Steilheit, der Mischsteilheit mit verschiedenen Zwischenfrequenzen und Bildung von Kombinationsfrequenzen dritter Ordnung, ist eine befriedigende quantitative Übereinstimmung der berechneten mit den gemessenen Werten festzustellen.

The r-f transconductance of a transistor depends on frequency and cannot be derived from the characteristic showing the collector d-c current as a function of the base-emitter d-c voltage. In this paper an empirical equation is given, which permits the calculation of the r-f transconductance as a function of the base-emitter d-c voltage U_{BE} up to a frequency nearly twice the α -cutoff frequency. The constants of this equation can be taken from the measurement of the transconductance as a function of U_{BE} at a single frequency in the vicinity of the α -cutoff frequency. The differentiation of this equation with respect to U_{BE} gives the higher derivatives of the transconductance needed for calculation of the conversion transconductance and the collector short circuit current at combination frequencies of higher order.

In the given basic equation (1) and the equations (6) and (7) derived from it the time constant of the high frequency transistor is subdivided into two parts:

τ_B for the base-emitter space and
 τ_C for the collector-base space.

Equation (7) shows that for calculation of the conversion transconductance the time constant τ_B is effective at the radio frequency (input frequency) and the time constant τ_C at the intermediate frequency (output frequency). By comparing measurements of the straight transconductance and conversion transconductance it is possible to separate the two time constants. The measurements show that for the high frequency alloy junction transistor the time constant τ_C is small compared with τ_B , and the conversion transconductance depends thus only slightly on the intermediate frequency. The drift transistor differs from the alloy junction transistor in this respect. For the drift transistor the time constant τ_C is larger than τ_B , and so the conversion transconductance depends strongly on the intermediate frequency. The equations of this paper, which are derived from the empirically found basic equation, were experimentally checked for the r-f alloy junction transistor AF 101 and the drift transistor OC 615. For the cases considered, viz. the transconductance, the conversion transconductance and the formation of third-order combination frequencies, the calculated and measured values show a good quantitative agreement.

Einleitung

Zur Beurteilung eines Transistors in Verstärker- und Mischschaltungen ist, neben dem Ein- und Ausgangsleitwert, die Kenntnis des Betrages der Steil-

heit bzw. der Mischsteilheit des Transistors von wesentlicher Bedeutung.

Die Herleitung der Steilheit aus der Collectorgleichstrom-Basisemittergleichspannungs-Kennlinie

ist für hohe Frequenzen wegen des Basiswiderstandes und der Laufzeiten im Inneren des Transistors nicht mehr möglich. Eine Reihe von Veröffentlichungen, z. B. [1], [2], behandelt dieses Problem unter Verwendung der von L. J. GIACOLETTO [3] angegebenen Transistor-Ersatzschaltung. Mit Hilfe dieser Ersatzschaltung gibt W. MOORTGAT-PICK [2] u. a. Gleichungen zur Berechnung der Steilheit an. Die in diesen Gleichungen vorkommenden Größen sind vorwiegend von der Meßfrequenz und vom Arbeitspunkt abhängig. Diese Abhängigkeit ist bei einigen dieser Größen in [2] (z. B. dem Stromverstärkungsfaktor β oder der Zeitkonstante $\tau_L = w^2/6Dp$) nicht durch Gleichungen ausgedrückt, sondern durch Meßkurven angegeben. Deshalb muß die Steilheit für jeden Arbeitspunkt und jede Frequenz mit Hilfe der Meßwerte einzeln berechnet werden und ein Differenzieren der Steilheit nach der Basisemittergleichspannung, wie dies u. a. zur Berechnung der Mischsteilheit erforderlich ist, läßt sich praktisch nur noch graphisch durchführen.

Außerdem läßt sich der Gl. (67) in [2], die lautet $|S| \approx |Y_{21}| \approx$

$$\approx \frac{S_0}{\sqrt{\left(1 + \frac{r_b S_0}{\alpha'}\right)^2 + \omega^2 r_b^2 (3 S_0 \tau_L + C_e + C_c)^2}},$$

entnehmen, daß für hohe Frequenzen, bei denen die Ungleichung

$$\left(1 + \frac{r_b S_0}{\alpha'}\right)^2 \ll \omega_1^2 r_b^2 (3 S_0 \tau_L + C_e + C_c)^2$$

erfüllt ist, und bei einem so großen Emitterstrom, daß

$$3 S_0 \tau_L \gg C_e + C_c$$

ist, der Betrag der Steilheit einem Grenzwert zustrebt, der durch die Beziehung

$$|S|_{\max} \approx |Y_{21}|_{\omega \gg \omega_1} \approx \frac{1}{\omega_1^2 r_b^2 3 \tau_L}$$

gegeben ist.

Messungen an Drifttransistoren zeigen jedoch bei hohen Frequenzen ein Maximum des Betrages der Steilheit als Funktion des Emitterstromes. Wird der Emitterstrom weiter erhöht, dann nimmt der Betrag der Steilheit wieder ab (siehe Bild 5).

In den folgenden Ausführungen wird eine sowohl für HF-Legierungs- als auch für Drifttransistoren gültige, empirisch ermittelte Gleichung angegeben, die die HF-Steilheitskennlinie als Funktion der Basisemittergleichspannung U_{BE} und der Arbeitsfrequenz für das Anwendungsgebiet von HF-Flächentransistoren bis etwa zur doppelten α -cutoff-Frequenz zu berechnen gestattet, wenn die Kennlinie bei einer Frequenz, die in der Nähe der α -cutoff-Frequenz liegt, aus Messungen vorliegt.

Ferner kann mit Hilfe der 1. und 3. Ableitung dieser Gleichung die zu erwartende Mischsteilheit berechnet und mit der 2. Ableitung der Betrag des Collectorkurzschlußstromes, der mit der Kombinationsfrequenz 3. Ordnung behaftet ist, bestimmt werden. Letzteres ist für die Beurteilung des Transistors bezüglich Kreuzmodulation und dessen Einsatz in Verstärkern ohne oder nur mit geringer Vorselektion von Bedeutung.

1. Der Betrag der Steilheit

Unter dem Betrag der Steilheit $|S|$ versteht man den an den Anschlußelektroden des Transistors gemessenen Betrag aus der Differenz des Übertragungskurzschlußleitwertes vorwärts und des Übertragungsleitwertes rückwärts (siehe Bild 1):

$$|S| = |Y_{21} - Y_{12}| = \left| \left(\frac{I_2}{U_1} \right)_{U_2=0} - \left(\frac{I_1}{U_2} \right)_{U_1=0} \right|.$$

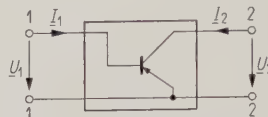


Bild 1. Transistor als allgemeiner Vierpol.

Nach eigenen Messungen an einer Vielzahl von Transistoren, sowohl HF-Legierungs- als auch Drifttransistoren, kann die Funktion des Betrages der Steilheit durch die empirisch ermittelte Gleichung

$$|S| \approx \frac{S_0}{\sqrt{1 + f/f_k}} \frac{1}{1 + 2\pi f(\tau_B + \tau_C)} \quad (1)$$

wiedergegeben werden. Darin ist

$S_0 = |S|_{f \rightarrow 0}$ die Steilheit bei tiefen Frequenzen,

f die Arbeitsfrequenz,

f_k eine Konstante mit der Dimension einer Frequenz; bei den untersuchten Transistortypen wurde $f_k \approx 100$ MHz gefunden,

τ_B die Zeitkonstante der Basisemitterstrecke,

τ_C die Zeitkonstante der Collectorbasisstrecke.

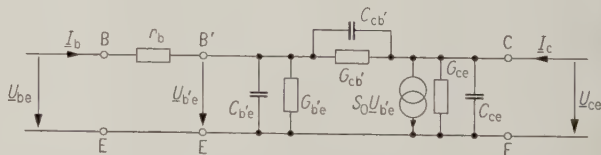


Bild 2. Transistorersatzschaltung nach L. J. GIACOLETTO.

Zur Erläuterung der Zeitkonstanten τ_B und τ_C dient das in Bild 2 gezeigte Ersatzschaltbild eines Flächentransistors nach L. J. GIACOLETTO [3].

Die Zeitkonstante τ_B stellt den Einfluß der Elemente r_b , $G_{b'e}$ und $C_{b'e}$ dar, wobei r_b der Basiswiderstand ist und der Leitwert $G_{b'e}$ den Anteil des über die Basis abfließenden Emitterflußstromes kennzeichnet, der vorwiegend durch Rekombination der Emitterdiode verursacht wird. Die Kapazität $C_{b'e}$ enthält die Sperrschicht- und die Emitterdiffusionskapazität.

Mit der Zeitkonstante τ_C wird der Einfluß der Elemente zwischen der inneren Basis B' und dem Collector C erfaßt. Der Leitwert $G_{cb'}$ setzt sich zusammen aus dem dynamischen Leitwert der Collectorsperrschicht und einer durch die EARLYsche Rückwirkung [4] verursachten Komponente. $C_{cb'}$ ist die Summe von Collectorsperrschicht- und Collectordiffusionskapazität. G_{ce} erfaßt den ebenfalls

durch die EARLYsche Rückwirkung [4] verursachten Ausgangsleitwert, und C_{ce} bildet die parallel erscheinende, im wesentlichen statische Ausgangskapazität. Ferner soll die Zeitkonstante τ_C noch die Einwirkung der Laufzeit der Ladungsträger im Basis-Collectorraum auf den Betrag des Collectorwechselstromes und damit auf die Steilheit S_0 berücksichtigen. S_0 wirkt im GIACOLETTOSchen Ersatzschaltbild als Einströmung $S_0 U_{b'e}$ zwischen Collector und Emitter.

Eine Erklärung für den Einfluß der Laufzeit der Ladungsträger im Basis-Collectorraum des Transistors auf den Betrag des Collectorwechselstromes möge durch folgende Überlegung gegeben sein:

Geht man von der Voraussetzung aus, daß der Transport der Ladungsträger zwischen Basisemittersperrschicht und Collectorbasissperrschicht auf verschiedenen Strompfaden geschieht, deren Länge unterschiedlich ist, so treffen die Ladungsträger, die die Basisemittersperrschicht zur gleichen Zeit verlassen, zu unterschiedlichen Zeiten an der Collectorsperrschicht ein. Dies hat bei hohen Arbeitsfrequenzen eine Abnahme der Wechselstromamplitude und damit eine Verminderung der Steilheit zur Folge. Die Ursache für die verschiedenen Längen der Strompfade kann darin gesehen werden, daß sowohl die Emitter- als auch die Collectorsperrschicht auf Grund von Emitter- und Collectorpille mehr oder weniger gewölbte Flächen sind. Ferner dürfte zur Abnahme der Amplitude des Collectorwechselstromes eine unterschiedliche Geschwindigkeit der Ladungsträger beitragen, die sich genauso wie eine unterschiedliche Länge der Strompfade auswirkt.

Die Zeitkonstanten τ_B und τ_C sind, wie Messungen zeigen, vom Collectorgleichstrom, d. h. vom Gleichstromarbeitspunkt des Transistors abhängig und damit Funktionen der Basisemittergleichspannung U_{BE} .

Ferner enthält die Gl. (1) noch den Korrekturfaktor $(1 + f/f_k)^{1/3}$, der vom Arbeitspunkt des Transistors unabhängig ist und bei sehr kleinen Emitterströmen den geringen Einfluß der Arbeitsfrequenz f auf den Betrag der Steilheit berücksichtigt (siehe Bild 5).

2. Die Bestimmung der Größen der Gleichung (1)

2.1. Die Steilheit für tiefe Frequenzen $S_0(f \rightarrow 0)$

Die Abhängigkeit des Collectorgleichstromes I_C von der Basisemittergleichspannung U_{BE} bei konstanter Collectoremmittergleichspannung U_{CE} wird durch Messung bestimmt. Trägt man I_C im logarithmischen Maßstab über U_{BE} auf, so erhält man bei HF-Legierungs- und Drifttransistoren bis zu Collectorgleichströmen $I_C \leq 5$ mA praktisch eine Gerade (siehe Bilder 3 und 14), so daß der Collectorgleichstrom als Funktion von U_{BE} geschrieben werden kann:

$$I_C \approx -i_0 \exp\left(\frac{-U_{BE}}{U_T'}\right). \quad (2)$$

Hierbei ist mit guter Näherung U_{BE} die Basisemittergleichspannung an den Anschlußelektroden

des Transistors, da bei den untersuchten HF-Legierungs- und Drifttransistoren der Stromverstärkungsfaktor in Emitterschaltung β_0 sehr groß und der Basiswiderstand r_b hinreichend klein sind, so daß bis zu Collectorgleichströmen $I_C \leq 5$ mA der Spannungsabfall an r_b vernachlässigt werden darf.

U_T' ist die reduzierte Temperaturspannung, die die Abhängigkeit des Stromverstärkungsfaktors β_0 vom Arbeitspunkt mitberücksichtigt. Die Steilheit $S_0(f \rightarrow 0)$ erhält man durch Bilden des Differentialquotienten

$$S_0 = \frac{dI_C}{dU_{BE}} \approx \frac{i_0}{U_T'} \exp\left(\frac{-U_{BE}}{U_T'}\right). \quad (3)$$

2.2. Die Summe der Zeitkonstanten τ_B und τ_C

Der Betrag der Steilheit $|S|$ wird bei einer Frequenz, die in der Nähe der α -cutoff-Frequenz der Transistortype liegt, in Abhängigkeit von der Basisemittergleichspannung U_{BE} gemessen. Das Meßverfahren ist in [5] beschrieben.

Für die Summe $\tau_B + \tau_C$ erhält man mit Gl. (1)

$$\tau_B + \tau_C \approx \frac{1}{2\pi f} \left(\frac{S_0}{|S| \sqrt{1 + f/f_k}} - 1 \right). \quad (4)$$

Da S_0 und $|S|$ verschieden vom eingestellten Arbeitspunkt sind und damit unterschiedlich von U_{BE} abhängen, muß auch die Summe $\tau_B + \tau_C$ eine Funktion von U_{BE} sein. Es läßt sich durch die in den Bildern 4 und 15 dargestellten Meßergebnisse zeigen, daß die Summe $\tau_B + \tau_C$ im exponentiellen Zusammenhang mit U_{BE} steht, so daß

$$\tau_B + \tau_C \approx (\tau_{B0} + \tau_{C0}) \exp\left(\frac{-U_{BE}}{U_\tau}\right) \quad (5)$$

geschrieben werden kann. U_τ ist eine vom Typ des Transistors abhängige Konstante mit der Dimension einer Spannung, die die Steigerung der exponentiellen Abhängigkeit der Zeitkonstanten τ_B und τ_C von U_{BE} angibt. U_τ ist von der Temperaturspannung verschieden.

2.3. Die Steilheit $|S|$ als Funktion von U_{BE}

Setzt man die Gl. (3) und (5) in die Gl. (1) ein, so erhält man

$$|S| \approx \frac{\frac{i_0}{U_T' \sqrt{1 + f/f_k}} \exp\left(\frac{-U_{BE}}{U_T'}\right)}{1 + 2\pi f(\tau_{B0} + \tau_{C0}) \exp\left(\frac{-U_{BE}}{U_\tau}\right)}. \quad (6)$$

3. Der Betrag der Mischsteilheit $|S_c|$

Die Mischsteilheit ist z. B. nach [6] definiert als das Verhältnis der Collectorkurzschlußstromänderung $d|I_{cZF}|$ mit der Zwischenfrequenz zur Basisemitterspannungsänderung $d|U_{beHF}|$ mit der Eingangshochfrequenz

$$|S_c| = \frac{d|I_{cZF}|}{d|U_{beHF}|}.$$

Diese Gleichung besagt, daß beim Differenzieren nicht von Gl. (6), die für die Geradeaussteilheit Gültigkeit hat, ausgegangen werden darf, sondern

diese Gleichung so zu modifizieren ist, daß die Zeitkonstanten τ_B und τ_C mit den jeweilig an ihnen wirksamen Frequenzen zur Geltung kommen. Dies bedeutet:

Die an den Eingang des Transistors bei Mischbetrieb angelegte, hochfrequente Eingangsspannung \underline{U}_{beHF} mit der Frequenz f_{HF} wird bis zur Steuerstrecke B'E des Transistors durch die Zeitkonstante τ_B beeinflusst. An der Steuerstrecke des Transistors erfolgt bei Vorhandensein einer Oszillatorspannung die Mischung und es entsteht unter anderem ein Strom mit der Zwischenfrequenz f_{ZF} . Für diesen Strom \underline{I}_{cZF} , der zwischen innerer Basis B' und dem Collector C fließt, ist die Zeitkonstante τ_C mit der Zwischenfrequenz f_{ZF} wirksam.

Für den Mischbetrieb ist somit die Zeitkonstante der Basisemitterstrecke τ_B mit der Hochfrequenz f_{HF} und die Zeitkonstante der Collectorbasisstrecke τ_C mit der Zwischenfrequenz f_{ZF} in Gl. (6) zu multiplizieren. Bezeichnet man die dahingehend abgewandelte Steilheit mit $|S_z|$, so geht Gl. (6) über in

$$|S_z| \approx \frac{i_0}{U_T' \sqrt{1 + f_{HF}/f_k}} \exp\left(\frac{-U_{BE}}{U_T'}\right) \cdot \frac{1}{1 + 2\pi f_{HF}(\tau_{B0} + \tau_{C0} \frac{f_{ZF}}{f_{HF}}) \exp\left(\frac{-U_{BE}}{U_T}\right)} \quad (7)$$

Durch vergleichende Messungen der Geradeaussteilheit $|S|$ und der Mischsteilheit $|S_c|$ bei verschiedenen Zwischenfrequenzen kann das Verhältnis τ_{B0}/τ_{C0} ermittelt werden. Für je einen typischen Vertreter der Drift- und HF-Legierungstransistorgruppe ergibt sich:

Drifttransistor OC 615 (Telefunken)

$$\tau_{B0}/\tau_{C0} \approx 1/2,$$

HF-Legierungstransistor AF 101 (Telefunken)

$$\tau_{B0}/\tau_{C0} \approx 4/1.$$

Man erkennt daraus in Verbindung mit Gl. (7), daß beim Drifttransistor, bei dem τ_{C0} doppelt so groß wie τ_{B0} ist, die Mischsteilheit deutlich von der gewählten Zwischenfrequenz abhängig ist. Beim HF-Legierungstransistor dagegen beträgt τ_{C0} nur 1/4 von τ_{B0} , so daß eine Abhängigkeit der Mischsteilheit von der Zwischenfrequenz kaum in Erscheinung tritt.

Für die Mischsteilheit $|S_c|$ unterscheidet man je nach Amplitude der Oszillatorwechselspannung den linearen und den nichtlinearen Fall.

3.1. Die Mischsteilheit $|S_c|$ für den linearen Fall

Die Mischsteilheit $|S_c|$ für den linearen Fall kann nach [7] wie folgt berechnet werden:

$$|S_c| \approx \frac{|S_z|'}{2} \underline{U}_2 + \frac{|S_z|'''}{16} (\underline{U}_1^2 \underline{U}_2 + \underline{U}_2^3). \quad (8)$$

Sie wird aus der 1. Ableitung ($|S_z|'$, Gl. (11)) und der 3. Ableitung ($|S_z|'''$, Gl. (13)) nach U_{BE} der modifizierten Steilheit $|S_z|$ nach Gl. (7) bestimmt. Hierbei ist zu beachten, daß die Eingangskennlinie des Transistors durch die Steuerwechselspannungen \underline{U}_1 und \underline{U}_2 nicht übersteuert wird. Da die Empfangsfrequenz-Steuerspannung \underline{U}_1 im allgemeinen

kleiner als die Oszillatorfrequenz-Steuerspannung \underline{U}_2 ist, gilt für den linearen Fall:

$$\begin{aligned} \text{Drifttransistor OC 615} & \quad \underline{U}_2 \leq 20 \text{ mV}, \\ \text{HF-Legierungstransistor AF 101} & \quad \underline{U}_2 \leq 40 \text{ mV}. \end{aligned}$$

In vielen Fällen ist der 2. Summand der Gl. (8) klein gegen den 1. Summanden, dann gilt in erster Näherung

$$|S_c| \approx \frac{|S_z|'}{2} \underline{U}_2. \quad (8a)$$

3.2. Die Mischsteilheit $|S_c|_{\max}$ für den nichtlinearen Fall

Bei voller Durchsteuerung der Transistoreingangskennlinie, d. h. bei Schaltfunktion des Transistors, wird in einer Arbeit von H. BENEKING [8] gezeigt, daß die Mischsteilheit

$$|S_c|_{\max} \approx \frac{1}{\pi} |S_{z,2}| \quad (9)$$

ist. Darin ist nach [8] für $|S_{z,2}|$ die modifizierte Steilheit $|S_z|$ einzusetzen, die zu dem doppelten Wert des mittleren Emitttergleichstromes gehört, der sich bei der Durchsteuerung einstellt.

$|S_c|_{\max}$ wird erreicht, wenn die Oszillatorfrequenz-Steuerspannung \underline{U}_2

für den Drifttransistor OC 615

$$\underline{U}_2 \geq 60 \text{ mV},$$

für den HF-Legierungstransistor AF 101

$$\underline{U}_2 \geq 120 \text{ mV}.$$

ist.

4. Betrag des für die Kombinationsfrequenzen

3. Ordnung wirksamen Collectorkurzschlußstromes

$$\underline{I}_{(2f_1 \pm f_2)}$$

Bei Vorhandensein der Summe zweier Wechselspannungen der Frequenz f_1 und f_2 an der Steuerstrecke des Transistors entsteht u. a. ein Collectorkurzschlußstrom mit der Kombinationsfrequenz 3. Ordnung $\underline{I}_{(2f_1 \pm f_2)}$.

Für den Betrag dieses Collectorkurzschlußstromes folgt nach [7], wenn für den Transistor die Frequenz f_1 nicht sehr von f_2 abweicht,

$$|\underline{I}_{(2f_1 \pm f_2)}| \approx \frac{|S|''}{8} \underline{U}_1^2 \underline{U}_2, \quad (10)$$

wobei $|S|''$ die 2. Ableitung der Steilheit $|S|$ nach U_{BE} entsprechend Gl. (12) ist. Diese Beziehung wird im Abschnitt 7 b benutzt.

5. Die Ableitungen der Steilheit $|S|$ bzw. $|S_z|$ nach U_{BE}

Faßt man aus Gl. (6) bzw. (7) zusammen:

$$A = \frac{i_0}{U_T' \sqrt{1 + f/f_k}} \quad \text{bzw.} \quad \frac{i_0}{U_T' \sqrt{1 + f_{HF}/f_k}},$$

$$B = 2\pi f(\tau_{B0} + \tau_{C0}) \quad \text{bzw.} \quad 2\pi f_{HF} \left(\tau_{B0} + \tau_{C0} \frac{f_{ZF}}{f_{HF}} \right),$$

so erhält man:

5.1. Die 1. Ableitung der Steilheit $|S|$

$$|S|' = \frac{d|S|}{dU_{BE}}, \quad |S|' = \frac{\frac{A}{U_T'} \exp\left(\frac{-U_{BE}}{U_T'}\right)}{1 + B \exp\left(\frac{-U_{BE}}{U_\tau}\right)} - \frac{\frac{A B}{U_\tau} \exp\left[-U_{BE}\left(\frac{1}{U_T'} + \frac{1}{U_\tau}\right)\right]}{\left[1 + B \exp\left(\frac{-U_{BE}}{U_\tau}\right)\right]^2}. \quad (11)$$

5.2. Die 2. Ableitung der Steilheit $|S|$

$$|S|'' = \frac{d^2|S|}{dU_{BE}^2}, \quad |S|'' = \frac{\frac{A}{U_T'^2} \exp\left(\frac{-U_{BE}}{U_T'}\right)}{1 + B \exp\left(\frac{-U_{BE}}{U_\tau}\right)} - \frac{\frac{2 A B}{U_T' U_\tau} \exp\left[-U_{BE}\left(\frac{1}{U_T'} + \frac{1}{U_\tau}\right)\right]}{\left[1 + B \exp\left(\frac{-U_{BE}}{U_\tau}\right)\right]^2} +$$

$$+ \frac{\frac{A}{U_\tau^2} \left\{ B^2 \exp\left[-U_{BE}\left(\frac{1}{U_T'} + \frac{2}{U_\tau}\right)\right] - B \exp\left[-U_{BE}\left(\frac{1}{U_T'} + \frac{1}{U_\tau}\right)\right] \right\}}{\left[1 + B \exp\left(\frac{-U_{BE}}{U_\tau}\right)\right]^3}. \quad (12)$$

5.3. Die 3. Ableitung der Steilheit $|S|$

$$|S|''' = \frac{d^3|S|}{dU_{BE}^3}, \quad |S|''' = \frac{\frac{A}{U_T'^3} \exp\left(\frac{-U_{BE}}{U_T'}\right)}{1 + B \exp\left(\frac{-U_{BE}}{U_\tau}\right)} - \frac{\frac{3 A B}{U_T'^2 U_\tau} \exp\left[-U_{BE}\left(\frac{1}{U_T'} + \frac{1}{U_\tau}\right)\right]}{\left[1 + B \exp\left(\frac{-U_{BE}}{U_\tau}\right)\right]^2} +$$

$$+ \frac{\frac{3 A}{U_T' U_\tau^2} \left\{ B^2 \exp\left[-U_{BE}\left(\frac{1}{U_T'} + \frac{2}{U_\tau}\right)\right] - B \exp\left[-U_{BE}\left(\frac{1}{U_T'} + \frac{1}{U_\tau}\right)\right] \right\}}{\left[1 + B \exp\left(\frac{-U_{BE}}{U_\tau}\right)\right]^3} -$$

$$- \frac{\frac{A}{U_\tau^3} \left\{ B^3 \exp\left[-U_{BE}\left(\frac{1}{U_T'} + \frac{3}{U_\tau}\right)\right] - 4 B^2 \exp\left[-U_{BE}\left(\frac{1}{U_T'} + \frac{2}{U_\tau}\right)\right] + B \exp\left[-U_{BE}\left(\frac{1}{U_T'} + \frac{1}{U_\tau}\right)\right] \right\}}{\left[1 + B \exp\left(\frac{-U_{BE}}{U_\tau}\right)\right]^4}. \quad (13)$$

6. Die Grundgrößen des Drifttransistors OC 615

Entsprechend den Ausführungen im Abschnitt 2 werden in diesem Abschnitt die Konstanten der Gl. (6) und (7) bestimmt. Für den Drifttransistor OC 615 wurden $I_C = f(U_{BE})$ und $|S| = f(U_{BE})$ bei 87 MHz gemessen. Die Meßergebnisse eines Mittelwerttransistors zeigt Bild 3. Man erhält aus $I_C = f(U_{BE})$ die Konstanten i_0 und U_T' sowie mit Gl. (3) $S_0 = f(U_{BE})$. Mit Gl. (4) kann die Summe $\tau_B + \tau_C$ der Zeitkonstanten als Funktion von U_{BE} ermittelt werden. S_0 und $\tau_B + \tau_C$ als Funktion von U_{BE} zeigt Bild 4. Die letzte Funktion liefert die Konstanten $\tau_{B0} + \tau_{C0}$ und U_τ . Mit dem Verhältnis τ_{B0}/τ_{C0} für den Drifttransistor, das durch vergleichende Messungen, wie im Abschnitt 3 ausgeführt, ermittelt wurde, ergibt sich

$$\tau_{B0} \approx \frac{1}{3} (\tau_{B0} + \tau_{C0}), \quad \tau_{C0} \approx \frac{2}{3} (\tau_{B0} + \tau_{C0}).$$

Damit lauten die Konstanten des Drifttransistors OC 615 bei $U_{CE} = -6 \text{ V}$, $T = 22^\circ \text{ C}$:

$$i_0 = 98 \cdot 10^{-9} \text{ A},$$

$$U_T' = 24,1 \text{ mV}, \quad U_\tau = 19 \text{ mV},$$

$$\tau_{B0} = 3,26 \cdot 10^{-15} \text{ s}, \quad \tau_{C0} = 6,52 \cdot 10^{-15} \text{ s}.$$

7. Überprüfung der angegebenen Gleichungen, angewandt auf den Drifttransistor OC 615

7.1. Überprüfung der Gleichungen (1) bzw. (6)

Die HF-Steilheit $|S|$ wurde nach Gl. (6) als Funktion von U_{BE} für die Frequenzen 1; 10; 30; 50; 87

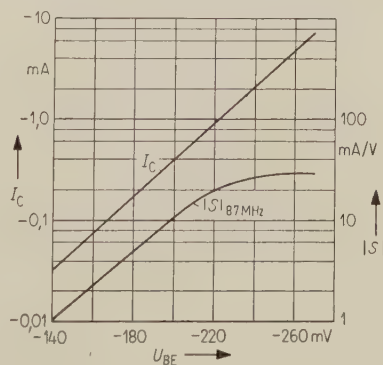


Bild 3. Gemessene Werte des Collectorgleichstromes I_C und des Betrages der HF-Steilheit $|S|$ für 87 MHz in Abhängigkeit von U_{BE} (Drifttransistor OC 615, $U_{CE} = -6 \text{ V}$, $T = 22^\circ \text{ C}$; $I_C \approx -i_0 \exp(-U_{BE}/U_T')$, $i_0 = 98 \cdot 10^{-9} \text{ A}$, $U_T' = 24,1 \text{ mV}$).

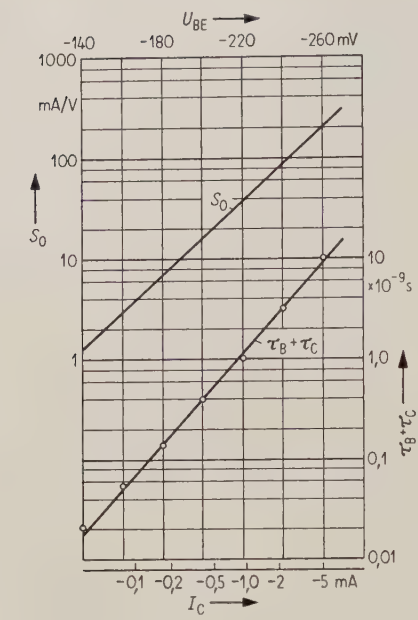


Bild 4. Berechnete Werte von S_0 nach Gl. (3) und $\tau_B + \tau_C$ nach Gl. (4). (Drifttransistor OC 615, $U_{CE} = -6$ V, $T = 22^\circ$ C; $S_0 \approx i_0 \exp(-U_{BE}/U_T)/U_T$, $i_0 = 98 \cdot 10^{-9}$ A, $U_T = 24,1$ mV; $\tau_B + \tau_C \approx (\tau_{B0} + \tau_{C0}) \exp(-U_{BE}/U_\tau)$, $\tau_{B0} + \tau_{C0} = 9,78 \cdot 10^{-15}$ s, $U_\tau = 19$ mV).

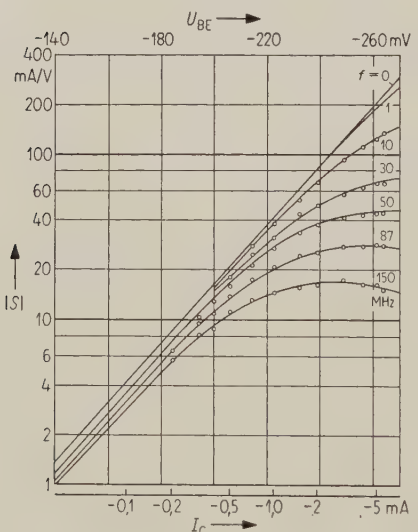


Bild 5. HF-Steilheit $|S|$ als Funktion von U_{BE} und der Frequenz f , berechnet nach Gl. (1) bzw. (6); Drifttransistor OC 615, $U_{CE} = -6$ V, $T = 22^\circ$ C;

$$|S| \approx \frac{S_0}{\sqrt[3]{1 + \frac{f}{100 \text{ MHz}}}} \frac{1}{1 + 2\pi f(\tau_B + \tau_C)};$$

o gemessene, — berechnete Werte.

und 150 MHz berechnet und mit den gemessenen Werten verglichen. Bild 5 zeigt die Ergebnisse. Eine gute Übereinstimmung für den gesamten Anwendungsfrequenzbereich des Transistors OC 615 kann festgestellt werden.

7.2. Überprüfung der Gleichungen (11), (12) und (13) in Verbindung mit Gleichung (6)

Für die Frequenz 87 MHz wurden die 1., 2. und 3. Ableitung der Steilheit $|S|$ als Funktion von U_{BE} berechnet und in den Bildern 6, 7 und 8 dargestellt.

a) Die Mischsteilheit $|S_c|$ wurde als Funktion von U_{BE} und der Oszillatorspannung U_2 gemessen,

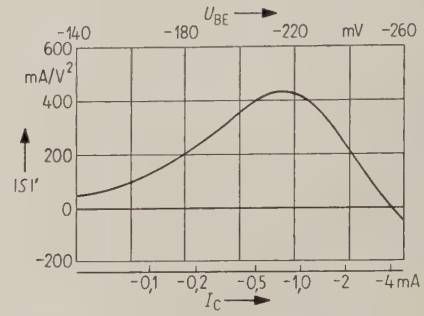


Bild 6. 1. Ableitung der Steilheit $|S|$ für $f_{ZF}/f_{HF} \approx 1$ nach Gl. (11) (Drifttransistor OC 615, $U_{CE} = -6$ V, $T = 22^\circ$ C, $f = 87$ MHz).

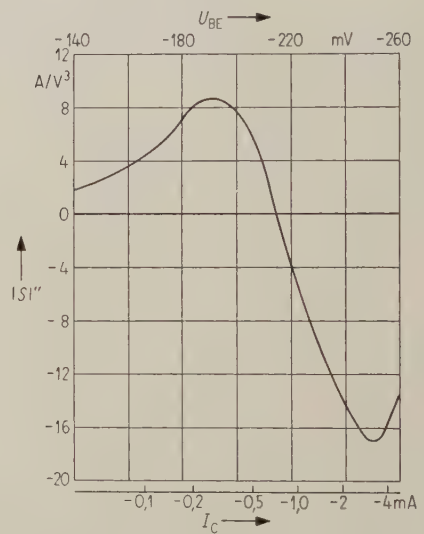


Bild 7. 2. Ableitung der Steilheit $|S|$ für $f_{ZF}/f_{HF} \approx 1$ nach Gl. (12) (Drifttransistor OC 615, $U_{CE} = -6$ V, $T = 22^\circ$ C, $f = 87$ MHz).

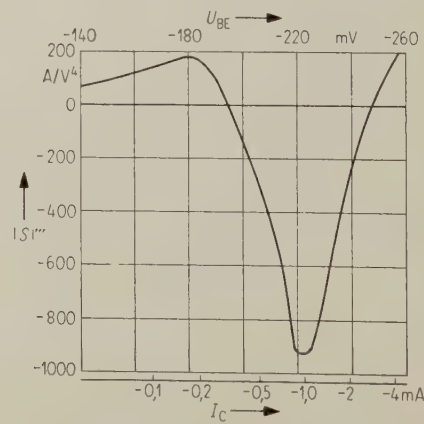


Bild 8. 3. Ableitung der Steilheit $|S|$ für $f_{ZF}/f_{HF} \approx 1$ nach Gl. (13) (Drifttransistor OC 615, $U_{CE} = -6$ V, $T = 22^\circ$ C, $f = 87$ MHz).

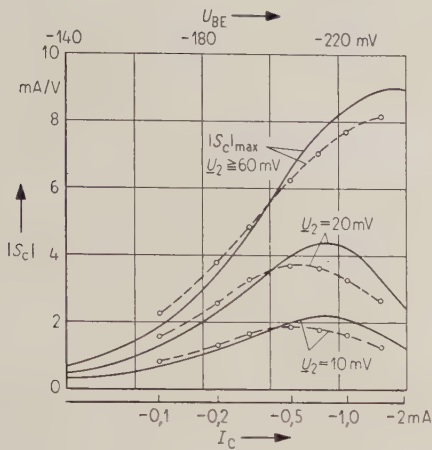


Bild 9. Mischsteilheit $|S_c|$ für $f_{ZF}/f_{HF} \approx 1$; für den linearen Fall $U_2 \leq 20$ mV nach Gl. (8a), für den nichtlinearen Fall $U_2 \geq 60$ mV nach Gl. (9) (Drifttransistor OC 615, $U_{CE} = -6$ V, $T = 22^\circ$ C; U_1 Eingangsspannung mit $f_{HF} = 93$ MHz, U_2 Oszillatorspannung mit $f_{osz} = 6$ MHz, $f_{ZF} = 87$ MHz); —○— gemessene, — berechnete Werte.

wobei die Zwischenfrequenz zu $f_{ZF} = 87$ MHz und die Eingangsfrequenz zu $f_{HF} = 93$ MHz gewählt wurde, d. h. $f_{ZF}/f_{HF} \approx 1$; damit geht Gl. (7) in Gl. (6) über. Die Oszillatorfrequenz betrug $f_{osz} = 6$ MHz.

Die nach Gl. (8a) und Gl. (9) berechneten und die durch Messung bestimmten Werte sind in Bild 9 gezeigt.

Für den linearen Fall, Gl. (8a), wird gute Übereinstimmung von Rechnung und Messung gefunden, dies trifft besonders bei $U_{BE} = -195$ mV zu, weil hier die 3. Ableitung, vergleiche Bild 6, eine Nullstelle hat, so daß die Näherungsgleichung (8a) der Grundgleichung (8) entspricht. Die Abweichungen der Werte nach Gl. (8a) unterhalb von $U_{BE} = -195$ mV betragen bis zu 15%, darüber bis zu 25%. Der Fehler wird bei Berücksichtigung des 2. Summanden in Gl. (8) kleiner. In der Praxis reicht die Berechnung der Mischsteilheit $|S_c|$ nach der Näherungsgleichung (8a) in den meisten Fällen aus.

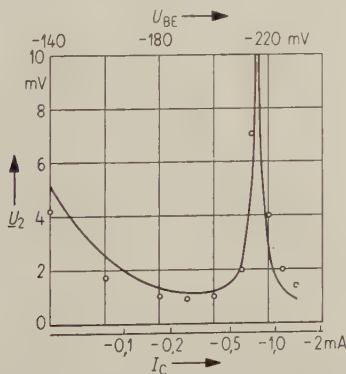


Bild 10. Steuerspannung U_2 als Funktion von U_{BE} für $|I_{(2f_1-f_2)}| = \text{const}$ nach Gl. (10) (Drifttransistor OC 615, $U_{CE} = -6$ V, $T = 22^\circ$ C; $|I_{(2f_1-f_2)}| = 0,715 \cdot 10^{-6}$ A = const, $2f_1 - f_2 = 87$ MHz, $U_1 = 25$ mV, $f_1 = 93,5$ MHz, $U_2 = f(U_{BE})$, $f_2 = 100$ MHz); ○ gemessene, — berechnete Werte.

Für den nichtlinearen Fall, Gl. (9), sind die maximalen Abweichungen der berechneten von der gemessenen Kurve kleiner als 20%. Der Fehler wird darauf zurückzuführen sein, daß Gl. (9) nach der Arbeit von H. BENEKING [8] eine rechteckförmige Umschaltspannung an der Steuerstrecke des Transistors voraussetzt, die wegen der in der Basiszone gespeicherten Minoritätsträger bei hohen Frequenzen nicht mehr voll realisierbar sein dürfte.

- b) Für einen konstanten Collectorkurzschlußstrom mit der Kombinationsfrequenz 3. Ordnung wurde die erforderliche Spannung U_2 bei Vorhandensein einer konstanten Spannung U_1 am Eingang des Transistors als Funktion von U_{BE} gemessen. Die Frequenz der Spannung U_1 war $f_1 = 93,5$ MHz, die der Spannung U_2 war $f_2 = 100$ MHz. Die Frequenz des Collectorkurzschlußstromes betrug $2f_1 - f_2 = 87$ MHz. Nach Gl. (10) erhält man für $|I_{(2f_1-f_2)}| = \text{const}$

$$U_2 = \frac{|I_{(2f_1-f_2)}| \cdot 8}{|S|'' U_1^2}.$$

Bild 10 zeigt die berechneten und die gemessenen Werte. Die Übereinstimmung ist gut. Die quadratische Abhängigkeit des Collectorkurzschlußstromes mit der Kombinationsfrequenz 3. Ordnung von der Eingangsspannung U_1 (f_1) und die lineare Abhängigkeit von der Eingangsspannung U_2 (f_2) wurden überprüft und bestätigt gefunden.

Eine nähere Betrachtung der Kurve in Bild 10 zeigt, daß für einen bestimmten Arbeitspunkt des Transistors die Eingangsspannung U_2 für konstanten Strom $I_{(2f_1-f_2)}$ sehr groß werden kann. Dies wurde bereits von M. AKGÜN und M. J. O. STRUTT [9] und H. LOTSCH [10] gezeigt. Es empfiehlt sich deshalb, diesen Arbeitspunkt des Transistors für Anfangsstufen zu wählen, um Kombinationsfrequenzen 3. Ordnung zu vermindern, die bei geringer Vorselektion und entsprechend großen Eingangssignalspannungen U_1 und U_2 im Transistor entstehen.

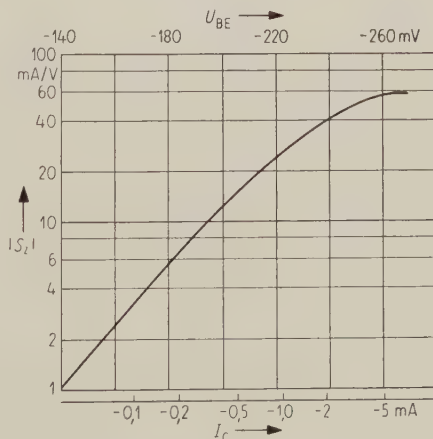


Bild 11. Steilheit $|S_z|$ für $f_{ZF}/f_{HF} \approx 0,115$ nach Gl. (7) (Drifttransistor OC 615, $U_{CE} = -6$ V, $T = 22^\circ$ C, $f_{HF} = 87$ MHz, $f_{ZF} = 10$ MHz, $f_{osz} = 97$ MHz).

7.3. Überprüfung der Gleichung (7)

Die Mischsteilheit $|S_c|$ wurde gemessen und mit den berechneten Werten, die sich aus Gl. (11) in Verbindung mit Gl. (7) und Gl. (8a) ergeben, überprüft. Hierbei war die Zwischenfrequenz $f_{ZF} = 10$ MHz, die Eingangsfrequenz $f_{HF} = 87$ MHz, die Oszillatorfrequenz $f_{osz} = 97$ MHz. Das Verhältnis ist $f_{ZF}/f_{HF} = 0,115$. Das Ergebnis der Rechnung nach Gl. (7) zeigt Bild 11 und die 1. Ableitung von $|S_z|$ nach U_{BE} mit Hilfe der Gl. (11) Bild 12.

Mit Benutzung von Gl. (8a) konnte die Mischsteilheit $|S_c|$ für den linearen Fall berechnet und in Bild 13 mit den gemessenen Werten verglichen werden. Im gleichen Bild sind die nach Gl. (9) gerechneten Werte $|S_c|_{\max}$ für den nichtlinearen Fall und die zugehörigen Meßwerte aufgetragen. Eine hinreichende Übereinstimmung kann auch hier festgestellt werden.

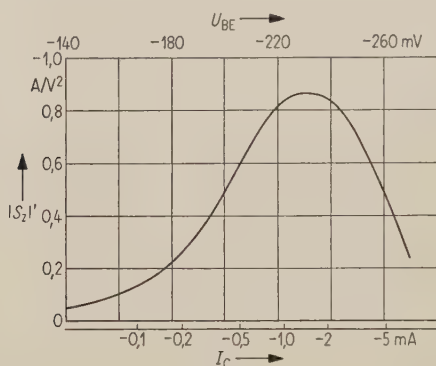


Bild 12. 1. Ableitung der Steilheit $|S_z|$ für $f_{ZF}/f_{HF} \approx 0,115$ nach Gl. (11) (Drifttransistor OC 615, $U_{CE} = -6$ V, $T = 22^\circ$ C, $f_{HF} = 87$ MHz, $f_{ZF} = 10$ MHz, $f_{osz} = 97$ MHz).

8. Die Grundgrößen des HF-Legierungstransistors AF 101

Nach den Ausführungen im Abschnitt 2 werden in diesem Abschnitt die Konstanten der Gl. (6) und (7) bestimmt. Für den HF-Legierungstransistor AF 101 wurde $I_C = f(U_{BE})$ und $|S| = f(U_{BE})$ bei 4 MHz gemessen. Die Meßergebnisse eines Mittelwerttransistors zeigt Bild 14.

Man erhält aus $I_C = f(U_{BE})$ die Konstanten i_0 und U_T' sowie mit Gl. (3) $S_0 = f(U_{BE})$. Mit Gl. (4) kann die Summe $\tau_B + \tau_C$ der Zeitkonstanten als Funktion von U_{BE} ermittelt werden. S_0 und $\tau_B + \tau_C$ sind als Funktion von U_{BE} in Bild 15 wiedergegeben. Die letzte Funktion liefert die Konstanten $\tau_{B0} + \tau_{C0}$ und U_T . Mit dem Verhältnis τ_{B0}/τ_{C0} für den HF-Legierungstransistor, das durch vergleichende Messungen, wie im Abschnitt 3 ausgeführt, ermittelt wurde, ergibt sich

$$\tau_{B0} \approx \frac{4}{5} (\tau_{B0} + \tau_{C0}), \quad \tau_{C0} \approx \frac{1}{5} (\tau_{B0} + \tau_{C0}).$$

Damit lauten die Konstanten des HF-Legierungstransistors AF 101 bei $U_{CE} = -6$ V, $T = 22^\circ$ C:

$$\begin{aligned} i_0 &= 450 \cdot 10^{-9} \text{ A}, \\ U_T' &= 25,9 \text{ mV}, & U_T &= 23,4 \text{ mV}, \\ \tau_{B0} &= 5,68 \cdot 10^{-12} \text{ s}, & \tau_{C0} &= 1,42 \cdot 10^{-12} \text{ s}. \end{aligned}$$

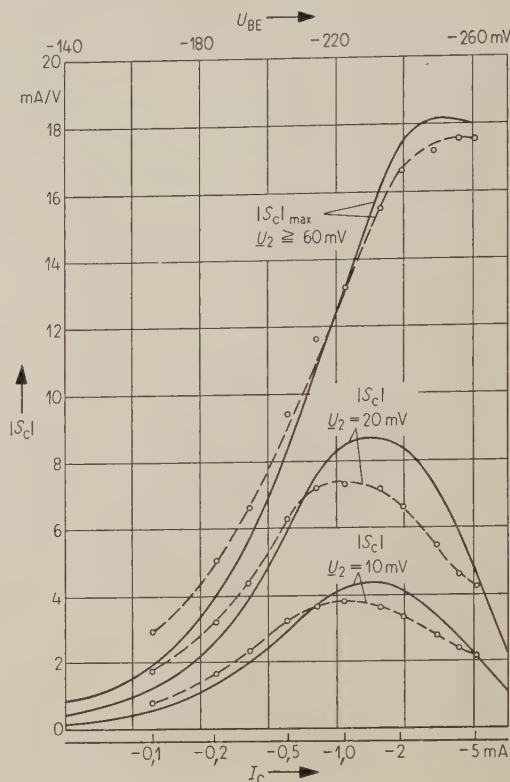


Bild 13. Mischsteilheit $|S_c|$ für $f_{ZF}/f_{HF} \approx 0,115$ für den linearen Fall $U_2 \leq 20$ mV nach Gl. (8a), für den nichtlinearen Fall $U_2 \geq 60$ mV nach Gl. (9) (Drifttransistor OC 615, $U_{CE} = -6$ V, $T = 22^\circ$ C, U_1 Eingangsspannung mit $f_{HF} = 87$ MHz, U_2 Oszillatorspannung mit $f_{osz} = 97$ MHz, $f_{ZF} = 10$ MHz); —○— gemessene, — berechnete Werte.

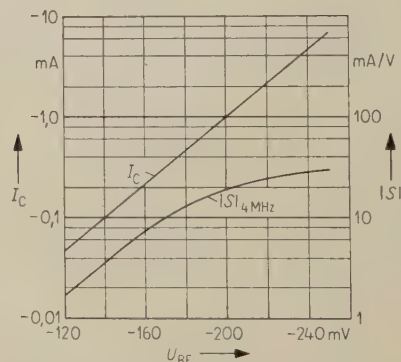


Bild 14. Gemessene Werte des Collectorgleichstromes I_C und des Betrages der HF-Steilheit $|S|$ für 4 MHz in Abhängigkeit von U_{BE} (HF-Legierungstransistor AF 101, $U_{CE} = -6$ V, $T = 22^\circ$ C; $I_C \approx -i_0 \exp(-U_{BE}/U_T')$, $i_0 = 450 \cdot 10^{-9}$ A, $U_T' = 25,9$ mV).

9. Überprüfung der angegebenen Gleichungen, angewandt auf den HF-Legierungstransistor AF 101

9.1. Überprüfung der Gleichungen (1) bzw. (6)

Die HF-Steilheit $|S|$ wurde nach Gl. (6) als Funktion von U_{BE} für die Frequenzen 0,5; 1; 2; 4 und 8 MHz berechnet und mit den gemessenen Werten

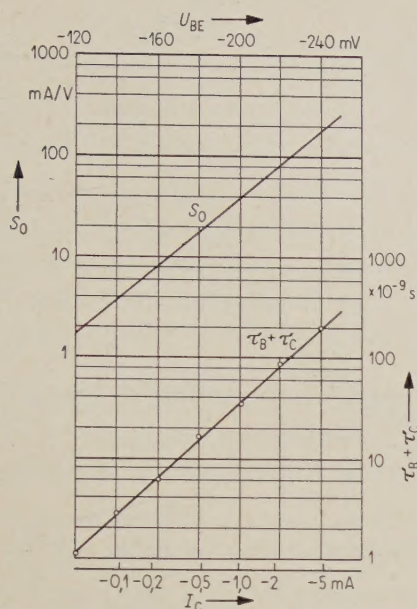


Bild 15. Berechnete Werte von S_0 nach Gl. (3) und $\tau_B + \tau_C$ nach Gl. (4) (HF-Legierungstransistor AF 101, $U_{CE} = -6$ V, $T = 22^\circ$ C; $S_0 \approx i_0 \exp(-U_{BE}/U_T)/U_T$, $i_0 = 450 \cdot 10^{-9}$ A, $U_T = 25,9$ mV; $\tau_B + \tau_C \approx (\tau_{B0} + \tau_{C0}) \exp(-U_{BE}/U_T)$, $\tau_{B0} + \tau_{C0} = 7,1 \cdot 10^{-12}$ s, $U_T = 23,4$ mV).

verglichen. Bild 16 zeigt die Ergebnisse. Für den gesamten Anwendungsbereich des HF-Legierungstransistors AF 101 besteht eine gute Übereinstimmung der berechneten mit den gemessenen Werten.

9.2. Überprüfung der Gleichung (11) in Verbindung mit Gleichung (6)

Für die Frequenz 4 MHz wurde die 1. Ableitung der Steilheit $|S|$ als Funktion von U_{BE} nach Gl. (11) berechnet und in Bild 17 dargestellt.

Zur Überprüfung wurde die Mischsteilheit $|S_c|$ als Funktion von U_{BE} und der Oszillatorspannung U_2 gemessen. Die Zwischenfrequenz wurde zu $f_{ZF} = 3,75$ MHz und die Eingangsfrequenz zu $f_{HF} = 4$ MHz gewählt, d. h. $f_{ZF}/f_{HF} \approx 1$; damit geht Gl. (7) in Gl. (6) über. Die Oszillatorfrequenz betrug $f_{Osz} = 0,25$ MHz.

Die nach Gl. (8a) berechneten und die durch Messung bestimmten Werte sind für Oszillatorspannungen U_2 von 10 mV und 30 mV in Bild 18 gezeigt. Eine gute Übereinstimmung kann für den linearen Fall festgestellt werden.

Bild 18. Mischsteilheit $|S_c|$ für $f_{ZF}/f_{HF} \approx 1$; für den linearen Fall $U_2 \leq 40$ mV nach Gl. (8a), für den nichtlinearen Fall $U_2 \geq 120$ mV nach Gl. (9) (HF-Legierungstransistor AF 101, $U_{CE} = -6$ V, $T = 22^\circ$ C, U_1 Eingangsspannung mit $f_{HF} = 4$ MHz, $U_2 =$ Oszillatorspannung mit $f_{Osz} = 0,25$ MHz, $f_{ZF} = 3,75$ MHz);
— o — gemessene, — berechnete Werte.

Die nach Gl. (9) berechneten und die durch Messung bestimmten Werte sind für den nichtlinearen Fall für $U_2 = 120$ mV ebenfalls in Bild 18 eingetragen. Auch hier ist die Übereinstimmung gut.

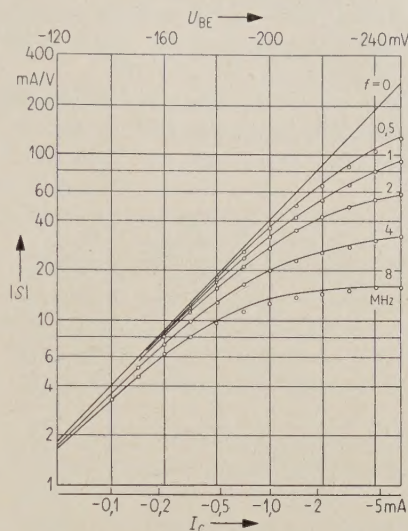


Bild 16. HF-Steilheit $|S|$ als Funktion von U_{BE} und der Frequenz f , berechnet nach Gl. (1) bzw. (6); HF-Legierungstransistor AF 101, $U_{CE} = -6$ V, $T = 22^\circ$ C;

$$|S| \approx \frac{S_0}{\sqrt[3]{1 + \frac{f}{100 \text{ MHz}}}} \frac{1}{1 + 2\pi f(\tau_B + \tau_C)};$$

o gemessene, — berechnete Werte.

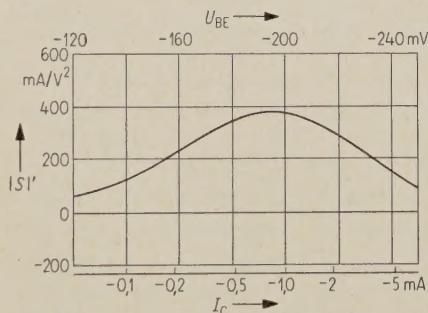
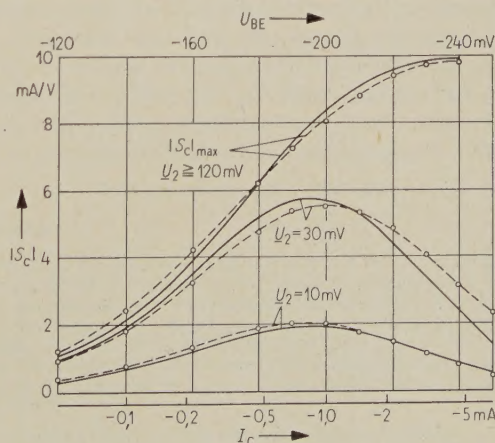


Bild 17. 1. Ableitung der Steilheit $|S|$ für $f_{ZF}/f_{HF} \approx 1$ nach Gl. (11) (HF-Legierungstransistor AF 101, $U_{CE} = -6$ V, $T = 22^\circ$ C, $f = 4$ MHz).



9.3. Überprüfung der Gleichung (7)

Die Mischsteilheit $|S_c|$ wurde gemessen und mit den berechneten Werten, die sich aus Gl. (11) in Verbindung mit Gl. (7) und Gl. (8a) ergeben, überprüft. Hierbei war die Zwischenfrequenz $f_{ZF} = 0,5$ MHz, die Eingangsfrequenz $f_{HF} = 4$ MHz, die Oszillatorfrequenz $f_{OSZ} = 3,5$ MHz.

Das Verhältnis ist $f_{ZF}/f_{HF} = 0,125$. Das Ergebnis der Rechnung für $|S_z|$ zeigt Bild 19 und die i. Ableitung von $|S_z|$ nach U_{BE} mit Hilfe der Gl. (11) Bild 20.

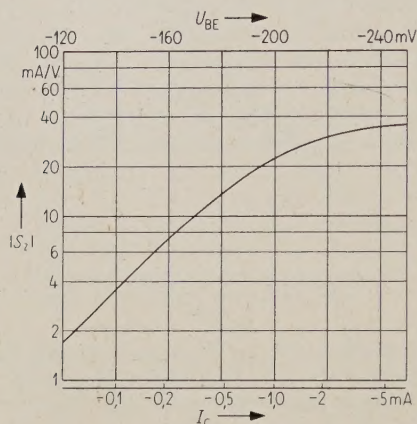


Bild 19. Steilheit $|S_z|$ für $f_{ZF}/f_{HF} = 0,125$ nach Gl. (7) (HF-Legierungstransistor AF 101, $U_{CE} = -6$ V, $T = 22^\circ$ C, $f_{HF} = 4$ MHz, $f_{ZF} = 0,5$ MHz, $f_{OSZ} = 3,5$ MHz).

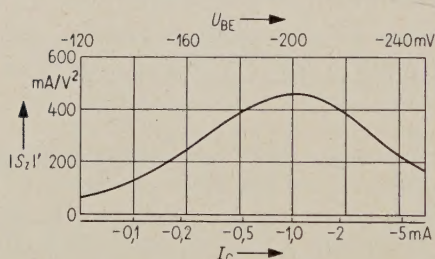


Bild 20. 1. Ableitung der Steilheit $|S_z|$ für $f_{ZF}/f_{HF} = 0,125$ nach Gl. (11) (HF-Legierungstransistor AF 101, $U_{CE} = -6$ V, $T = 22^\circ$ C, $f_{HF} = 4$ MHz, $f_{ZF} = 0,5$ MHz, $f_{OSZ} = 3,5$ MHz).

Die nach Gl. (8a) für den linearen Fall berechnete Mischsteilheit $|S_c|$ wird in Bild 21 mit den gemessenen Werten verglichen. Im gleichen Bild sind die für den nichtlinearen Fall nach Gl. (9) berechneten Werte für $|S_c|_{\max}$ und die zugehörigen Meßwerte aufgetragen. Die Übereinstimmung zwischen Messung und Rechnung ist für beide Fälle befriedigend.

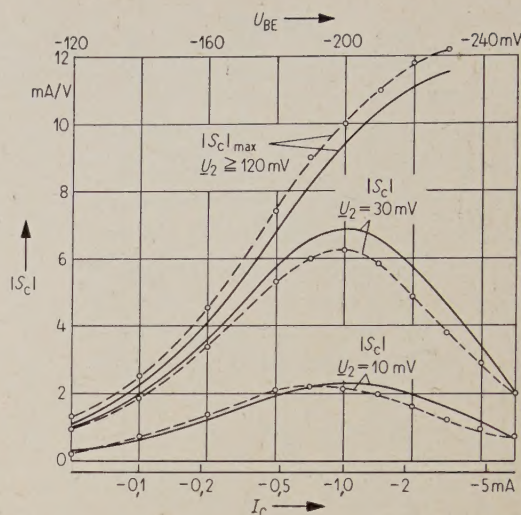


Bild 21. Mischsteilheit $|S_c|$ für $f_{ZF}/f_{HF} = 0,125$; für den linearen Fall $U_2 \leq 40$ mV nach Gl. (8a), für den nichtlinearen Fall $U_2 \geq 120$ mV nach Gl. (9) (HF-Legierungstransistor AF 101, $U_{CE} = -6$ V, $T = 22^\circ$ C, U_1 Eingangsspannung mit $f_{HF} = 4$ MHz, U_2 Oszillatorspannung mit $f_{OSZ} = 3,5$ MHz, $f_{ZF} = 0,5$ MHz); —○— gemessene, — berechnete Werte.

Den Herren Prof. Dr. Dr. M. J. O. STRUTT und Dipl.-Ing. J. S. VOGEL, E.T.H. Zürich, Prof. Dr. H. ROTHE, T.H. Karlsruhe, Dr. W. DAHLKE und Dr. J. SCHUBERT, Telefunken G.m.b.H., Ulm, danke ich für zahlreiche Diskussionen. Meinen Mitarbeitern, den Herren H. RINDERLE und Dipl.-Ing. H. TROGUS, die mir bei den umfangreichen Messungen und dem Aufbau der Meßgeräte behilflich waren, sei an dieser Stelle ebenfalls gedankt.

Schrifttum

- [1] MUELLER, C. W. und PANKOVE, J. I., A p-n-p triode alloy-junction transistor for radio-frequency amplification. Proc. Inst. Radio Engrs. **42** [1954], 386–391.
- [2] MOORTGAT-PICK, W., Ersatzschaltbild und Verstärkungseigenschaften des Flächentransistors. A.E.Ü. **13** [1959], 33–48, 82–89.
- [3] GIACOLETTO, L. J., Study of p-n-p-alloy-junction transistor from dc through medium frequencies. RCA Rev. **15** [1954], 506–562.
- [4] EARLY, J. M., Effect of space-charge layer widening in junction transistors. Proc. Inst. Radio Engrs. **40** [1952], 1401–1406.
- [5] MINNER, W., Die Vierpolkenngrößen des HF-Transistors OC 615 im Frequenzbereich von 30 bis 150 MHz. Telefunken Röhren-Mitt. Nr. 590 147 [1959].
- [6] ROTHE, H. und KLEEN, W., Elektronenröhren als Anfangsstufenverstärker. Akad. Verlagsges., Leipzig 1953, Kap. XIII, § 2, S. 205.
- [7] ROTHE, H. und KLEEN, W., Elektronenröhren als Anfangsstufenverstärker. Akad. Verlagsges., Leipzig 1953, Kap. IV, § 2, S. 29.
- [8] BENEKING, H., Ein Transistor-Mischer-Ersatzschaltbild. A.E.Ü. **13** [1959], 313–319.
- [9] ARGÜN, M. und STRUTT, M. J. O., Nichtlineare Verzerrungen einschließlich Kreuzmodulation in Hochfrequenz-Transistorstufen. A.E.Ü. **13** [1959], 227–242.
- [10] LOTSCH, H., Übersicht über die nichtlinearen Verzerrungen in Transistorstufen, einschließlich der Kreuzmodulation. A.E.Ü. **14** [1960], 204–216.

Mitarbeiter dieses Heftes:

Dr. W. Becker, Institut für Ionosphärenforschung, Lindau ü. b. Northeim/Hann.; Dr. H.-J. Butterweck, Institut für Hochfrequenztechnik der T. H., Aachen, Templergraben 55; Dipl.-Ing. R. Ebhardt, i. H. Siemens & Halske AG, München 8, Balanstraße 73; Prof. Dr. G. Eckart, Institut für angew. Physik und Elektrotechnik der Universität, Saarbrücken; Dr. R. Elsner, Institut für Fernmelde- und Hochfrequenztechnik der T. H., Braunschweig, Mühlenpfordtstraße 23; Dr. E. Groschwitz und Dr. E. Hofmeister, i. H. Siemens & Halske AG, München 8, Balanstraße 73; Dipl.-Ing. W. Minner, i. H. Telefunken GmbH, Ulm (Donau), Söflinger Straße 100; Prof. Dr. K. Rawer, Ionosphären-Institut, Breisach (Rhein); Dr. E. Schwartz, Rogowski-Institut für Elektrotechnik der T. H., Aachen, Templergraben 55; Dr. K. Sixtus, i. H. AEG-Forschungsinstitut, Frankfurt a. M.-Niederrad, Goldsteinstraße 235; Prof. Dr. Dr. M. J. O. Strutt und Dipl.-Ing. J. S. Vogel, Institut für höhere Elektrotechnik der E.T.H., Zürich 7, Gloriastraße 35.

Schriftleitung:

Dr. phil. J. Piesch,

Wien IV/50, Karlsplatz 13, Fernruf U 46530

Dr.-Ing. J. Schunack,

Berlin-Lichterfelde 1, Drakestraße 1a, Fernruf 732261

Prof. Dr.-Ing. E. h. Dr. techn. M. Strutt,

Zürich 7, Gloriastraße 35, Fernruf 327330

Oberstudiendirektor Dr. phil. A. Thoma,

Fulda, Heinrich-von-Bibra-Platz 1,

Fernruf (vorm.) 2962, (nachm.) 4428

Geschäftsführende Redaktion und Anzeigenverwaltung:

Ing. Friedrich Rühmann

(17a) Karlsruhe-Durlach, Machstraße 4, Fernruf 42854

Alle Zuschriften werden entweder an die Schriftleitung oder an die Geschäftsführende Redaktion erbeten.

Geschäftliche Hinweise:

Der Bezug des Archivs der elektrischen Übertragung durch Verlag, Post und Buchhandel ist im Abonnement zum Preis von DM 20,— für das Vierteljahr möglich. Einzelhefte können zum Preis von DM 7,— geliefert werden. Anzeigenaufträge und -anfragen bitten wir an die Anzeigenverwaltung der Zeitschrift zu richten.

S. HIRZEL VERLAG · STUTTGART

Postscheckkonto Stuttgart 64383

LAUFFELDRÖHREN

Von Prof. Dr. W. Kleen
und Dr. K. Pöschl

X, 192 Seiten mit 127 Bildern
16 cm × 23 cm, Ganzleinen DM 28,—

Teil 2 der

EINFÜHRUNG IN DIE MIKROWELLEN-ELEKTRONIK

Von Prof. Dr. Werner Kleen

Inhaltsübersicht

A. Wanderfeldröhre

1. Qualitative Übersicht
2. Lineare Vorgänge
3. Rauschen
4. Nichtlineare Vorgänge
5. Bemessung, Daten und Anwendungen von Wanderfeldröhren

B. Sonstige Lauffeldröhren

6. Rückwärtswellenröhre
 7. Elektronenwellenröhre
 8. Lauffeldröhre mit Widerstandsschicht
 9. Spezielle Formen von Lauffeldröhren und verwandte Effekte
- Anhänge

S. HIRZEL VERLAG · STUTTGART

VOM WESEN DER NACHRICHT

VON PROF. DR. G. MEGLA

etwa 190 Seiten mit etwa 70 Bildern,
mehreren Tafeln und Tabellen. Ganzleinen etwa DM 23,—

Erscheint im November 1960

1. Kapitel. Die Nachricht

- 1.1 Vorbetrachtung — 1.2 Die Symbole der Nachricht —
1.2.1 Die Symbolarten — 1.2.2 Die Empfangsbereitschaft — 1.3 Die unbewußte Nachrichtenübertragung — 1.4 Die bewußte Nachrichtenübertragung — 1.5 Der Weg einer Nachricht — 1.6 Die transportfähige Nachricht mit langer Übertragungszeit — 1.7 Die transportfähige Nachricht mit kurzer Übertragungszeit

2. Kapitel. Die Nachrichtenverbindungen

- 2.1 Die Träger der Nachrichtensymbole — 2.1.1 Das Signal als Träger des Symbols — 2.1.2 Mechanische Träger — 2.1.3 Der elektrische Strom als Symbolträger — 2.1.4 Die akustische Trägerschwingung — 2.1.5 Die elektromagnetische Trägerschwingung — 2.2 Die Teilstrecken einer Verbindung — 2.3 Optische und akustische Verbindungen — 2.4 Interplanetarische Verbindungen

3. Kapitel. Die elektrische Nachrichtenübertragung

- 3.1 Verfahren der elektrischen Nachrichtenübertragung — 3.2 Der Informationsgehalt der elektrischen Nachrichtenübertragung — 3.3 Neue Anwendungsgebiete

4. Kapitel. Die unmittelbare Nachrichtenübertragung

- 4.1 Vorbetrachtung — 4.2 Die nachrichtentechnischen Eigenschaften der Sinnesorgane — 4.3 Sinnesorgane und Nachrichtenverbindungen der Tiere — 4.4 Die Nachrichtenübertragung im Nervensystem

Literaturverzeichnis — Sachwortverzeichnis

Die modernen Nachrichtensysteme sind wie der Mensch in der Lage, Wahrnehmungen nicht nur aufzunehmen, sondern auch zu konzentrieren, zu speichern, zu verarbeiten und darüber hinaus ohne menschliche Hilfe logische automatisch überwachte Handlungen auszulösen. Aber auch mit diesen hochqualifizierten technischen Mitteln sind wir noch weit entfernt von Güte, Vielfalt und Zuverlässigkeit der nachrichtentechnischen Organe, wie sie das höchstentwickelte Lebewesen, der Mensch, besitzt.

Da die überaus schnelle Entwicklung der Technik die Gefahr birgt, die großen Zusammenhänge aus dem Auge zu verlieren, wird in diesem Buch der Versuch unternommen, Sinn und Wesen der Nachrichtenübertragung in einer möglichst umfassenden Gesamtschau darzustellen, deren Kenntnis es erleichtern wird, den richtigen Weg zu gehen und die gegebenen Möglichkeiten sinnvoll anzuwenden.

S. HIRZEL VERLAG · STUTTGART

